



# Réduction de la consommation statique des circuits intégrés en technologie SOI 65 nm partiellement désertée

Julien Le Coz Le Coz

## ► To cite this version:

Julien Le Coz Le Coz. Réduction de la consommation statique des circuits intégrés en technologie SOI 65 nm partiellement désertée. Autre. Université de Grenoble, 2011. Français. NNT : 2011GRENT076 . tel-00764400

**HAL Id: tel-00764400**

**<https://theses.hal.science/tel-00764400>**

Submitted on 13 Dec 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Micro et Nano Electronique**

Arrêté ministériel : 7 août 2006

Présentée par

**Julien LE COZ**

Thèse dirigée par **Marc BELLEVILLE** et  
codirigée par **Philippe FLATRESSE** et **Alexandre VALENTIAN**

préparée au sein du **Laboratoire CEA LETI**  
dans **l'École Doctorale Electronique Electrotechnique**  
**Automatique et Traitement du Signal (Grenoble-INP)**

# Réduction de la consommation statique des circuits intégrés en technologie SOI 65nm partiellement désertée

Thèse soutenue publiquement le **24 novembre 2011**,  
devant le jury composé de :

**M. Jean Michel FOURNIER**

Professeur à Grenoble-INP, IMEP-LAHC, Président

**M. Amara AMARA**

Professeur à l'ISEP, Rapporteur

**M. Jean Didier LEGAT**

Professeur à l'UCL, Rapporteur

**M. Christian FIGUET**

Professeur à l'EPFL, CSEM, Membre

**M. Philippe MAURINE**

Maitre de conférences, Université de Montpellier 2, Membre

**M. Marc BELLEVILLE**

Directeur de recherche au CEA LETI, Directeur de thèse et Membre

**M. Philippe FLATRESSE**

Chef de projet chez STMicroelectronics, Membre

**M. Alexandre VALENTIAN**

Ingénieur chercheur au CEA LETI, Membre





« A Seu João »



# Remerciements

Ces trois années de thèses ont été pour moi enrichissantes, professionnellement et humainement. Mes remerciements vont tout d'abord, et naturellement, à mon encadrant industriel, Philippe Flatresse. Je remercie Philippe de la confiance qu'il m'a accordée et de m'avoir fait profiter de son expertise technique en technologie SOI. Philippe a fortement participé à l'avancement des travaux de cette thèse.

Je remercie également mon employeur STMicroelectronics, et tout particulièrement Robin Wilson. Robin a été mon premier contact industriel alors que j'étais étudiant à Polytech'Montpellier ; son intérêt pour mon travail à cette époque a fortement contribué à mon embauche. Je pense également à Sylvain Engels, membre de l'équipe et expert en conception numérique, qui a su me livrer une multitude de conseils techniques et pour qui j'ai une grande admiration et amitié.

Cette thèse n'aurait pas eu lieu, sans l'appui du Commissariat Energie Atomique (CEA) et de son laboratoire d'électronique et technologie de l'information (LETI). Je remercie mon directeur de thèse, Marc Belleville, directeur de recherche au LETI, pour ses conseils pertinents et l'orientation claire qu'il a su donner à mes travaux. Il en va de même pour Alexandre Valentian, ingénieur chercheur au LETI. Alexandre fut mon encadrant de laboratoire ; son expertise en SOI et son talent de rédaction m'ont permis d'améliorer les miens.

Mes remerciements s'adressent également à l'ensemble du jury : au Professeur Jean Michel Fournier, pour avoir bien voulu accepter de présider celui-ci, aux Professeurs Amara Amara et Jean Didier Legat pour avoir accepté la lourde tâche d'en être les rapporteurs ainsi qu'aux autres membres, le Professeur Christian Piguet et le Docteur Philippe Maurine. Philippe, qui fut l'un de mes enseignants en école d'ingénieur ; il a ma plus respectueuse reconnaissance pour m'avoir donné goût à la microélectronique.

Au cours de ces trois années, à STMicroelectronics, j'ai eu la chance de rencontrer un grand nombre de personnes qui m'ont toutes directement ou indirectement permis d'avancer dans mes travaux, par leurs connaissances techniques, leurs disponibilités et leurs amitiés. J'aimerais d'abord remercier les collègues qui me sont les plus proches à STMicroelectronics, Alain Tournier, Carmelo D'Agostino, Thomas Le Huche, pour leur complicité et l'atmosphère favorable au travail qu'ils ont su créer dans cette équipe. Je remercie aussi ceux avec qui j'ai eu la chance de travailler, ainsi que les différentes personnes qui ont su

m'apporter des réponses. Je remercie Frank Gardic pour la conception du LDPC ; Philippe Galy pour ses cours de physique des semi-conducteurs ; Pascal Urard de son soutien pour ce qui fut ma plus belle expérience de doctorant : publier et présenter à l'ISSCC ; Damien Croain pour les mesures de circuits en PD-SOI et Sylvain Clerc pour l'optimisation de la bascule de rétention en PD-SOI. Je remercie d'autres collègues, autrefois doctorants, comme Fady Abouzeid, Nicolas Hebert, Lahcen Hamouche pour les bons moments passés à leur côté et également ceux qui sont devenus au fil des années de très bons amis, Dimitri Soussan, Thomas Benoist, Vincent Joseph. Mes remerciements vont aussi à Frédéric Bœuf, Vincent Barral, Benoist Lasbouygues, Andréia Cathelin et tous ceux qui ont décidé de tenter l'aventure ailleurs comme Luis Rolindez, Nicolas L'Hostis et Christophe Entringer.

Coté LETI, je remercie en priorité Christine Raynaud, pour son soutien et la volonté qu'elle avait de promouvoir la technologie PD-SOI. Elle m'a permis de connaître mieux cette technologie. Je remercie également Jean-Philippe Noel, co-thésard, Bastien Giraud et Olivier Thomas pour le travail en commun que nous avons eu en technologie FD-SOI.

Je tiens à remercier aussi Guillaume Cognard et Marc Esmerit de Dolphin Intégration. Mes pensées vont aussi à l'école d'ingénieur Polytech'Montpellier, le personnel enseignant et les anciens camarades de promotions MEA.

Ils m'ont indirectement permis d'avancer, je parle de mes amis, les plus proches et les autres, Lionel Ribeiro, Alexandre Bogdanski et ceux que je considère comme mes « fistons » Morgan Bourrée et Guillaume Roquet et le « petit frère » Tigrane Kaskarian ainsi que tous mes amis de la salle de sport « Univers Physique », lieu qui m'a permis de m'évader quelques heures par semaine.

Je remercie ma famille et ma belle-famille, mes parents, mon frère Anthony et principalement ma compagne, Marine. Je te remercie Marine pour ton soutien et ta tendresse, le temps que tu as passé à corriger ce manuscrit et surtout de m'avoir suivi dans cette aventure.

L'émotion me fait certainement oublier des personnes, sachez que mes remerciements vous sont aussi adressés.







<b>Remerciements</b>	<b>5</b>
<b>➤ Introduction</b>	<b>13</b>
<b>1. Les technologies SOI</b>	<b>19</b>
<b>1.1. Substrat et Transistor</b>	<b>20</b>
1.1.1. Substrat	20
1.1.2. Transistor	20
1.1.3. Pourquoi le SOI ?	21
1.1.4. Transistors PD-SOI et FD-SOI	21
<b>1.2. Comportement électrique des transistors PD-SOI</b>	<b>23</b>
1.2.1. Réduction des capacités de jonction	24
1.2.2. L'effet de body flottant	24
1.2.2.1. L'accumulation de charges dans le body flottant	26
1.2.2.2. Le couplage capacitif	26
1.2.2.3. L'effet d'histoire	27
1.2.3. Les courants du MOSFET	28
1.2.3.1. Courant actif	29
1.2.3.2. Courants de fuite	31
1.2.4. Comportement en température	35
1.2.5. Comparaison entre les technologies BULK et PD-SOI	35
1.2.5.1. Portage direct d'une mémoire SRAM	36
1.2.5.2. Portage direct d'un circuit de moyenne complexité	37
1.2.5.3. Consommation Statique	38
<b>1.3. Transistor avec prise body</b>	<b>39</b>
1.3.1. Inconvénient du transistor à body flottant (FB)	39
1.3.2. Le transistor à body contacté	40
1.3.3. Le transistor à body contacté connecté en DTMOs	42
<b>1.4. Conclusion</b>	<b>43</b>
<b>2. Les techniques de réduction de la consommation</b>	<b>49</b>
<b>2.1. Puissance dissipée dans les Circuits intégrés numériques</b>	<b>50</b>
2.1.1. Puissance dynamique et statique	50
2.1.2. Comparaison des besoins des applications fixes et portables	51
2.1.3. Evolution avec les technologies	51
<b>2.2. Techniques de réduction de la puissance dynamique</b>	<b>53</b>
2.2.1. La puissance dynamique	53
2.2.2. Réduction de la puissance dynamique	54
2.2.2.1. Réduction de l'alimentation VDD	54
2.2.2.2. Changement de fréquence	55
2.2.2.3. Optimisation de l'architecture	56
2.2.2.4. Contrôle des données et de l'horloge	56
2.2.2.5. Point de fonctionnement optimal	57

<b>2.3.</b>	<b>Techniques de réduction de la puissance statique</b>	<b>58</b>
2.3.1.	La puissance statique	58
2.3.2.	Réduction de la puissance statique	59
2.3.2.1.	Réduction de l'alimentation VDD	59
2.3.2.2.	Empilement de transistors	60
2.3.2.3.	Vecteur optimal	62
2.3.2.4.	Longueur de grille	62
2.3.2.5.	Polarisation de Substrat	62
2.3.2.6.	Polarisation Dynamique du body	65
2.3.2.7.	Conception Multi-Vt	68
<b>2.4.</b>	<b>Techniques basées sur des interrupteurs de puissance</b>	<b>70</b>
2.4.1.	Principe de base	70
2.4.2.	Dimensionnement des interrupteurs	71
2.4.3.	Contrôle et réveil	72
2.4.4.	Implémentation	73
2.4.5.	Les Bascules de rétention	73
<b>2.5.</b>	<b>Conclusion</b>	<b>74</b>
<b>3.</b>	<b><i>Optimisation des interrupteurs de puissance en 65nm PD-SOI</i></b>	<b>79</b>
<b>3.1.</b>	<b>Evolution des contraintes technologiques et problèmes spécifiques liés au PD-SOI</b>	<b>80</b>
3.1.1.	Besoin d'interrupteurs de puissance en 65nm	80
3.1.2.	Contexte de la technologie 0,13µm PD-SOI	81
3.1.3.	Comparaison des technologies PD-SOI et BULK en 65nm	81
3.1.4.	Contrainte supplémentaire apportée par l'effet de body flottant sur les interrupteurs de puissance	83
<b>3.2.</b>	<b>Conception et optimisation des interrupteurs de puissance, selon l'état de l'art</b>	<b>83</b>
3.2.1.	Introduction des paramètres électriques à optimiser : Ron et Roff	83
3.2.2.	Optimisation des solutions proposées dans l'état de l'art	85
3.2.2.1.	Description des solutions étudiées	85
3.2.2.2.	Choix de la structure pertinente	87
3.2.2.3.	Optimisation utilisant une polarisation du body des interrupteurs de puissance	90
3.2.3.	Optimisation de la surface d'implémentation	91
3.2.4.	Capacités de découplage	93
<b>3.3.</b>	<b>Proposition d'optimisation du Ron de l'interrupteur de puissance BC</b>	<b>97</b>
3.3.1.	Présentation du montage Auto-DTMOS proposé	97
3.3.2.	Implémentation du montage Auto-DTMOS en PD-SOI 65nm	100
3.3.3.	Développement de la bibliothèque associée au montage Auto-DTMOS et Réalisation du circuit de test	102
3.3.4.	Mesures et comparaisons des interrupteurs de puissance FB, BC et Auto-DTMOS	105
<b>3.4.</b>	<b>Conclusion sur l'Auto-DTMOS</b>	<b>108</b>
<b>4.</b>	<b><i>Application à un circuit LDPC</i></b>	<b>111</b>
<b>4.1.</b>	<b>Description du circuit LDPC</b>	<b>112</b>
<b>4.2.</b>	<b>Stratégie d'insertion des interrupteurs de puissance pour circuit LDPC</b>	<b>115</b>
4.2.1.	Conception d'un circuit	115
4.2.2.	Insertion des interrupteurs de puissance en anneau	116
4.2.3.	Dimensionnement des interrupteurs de puissance	118

<b>4.3. Comparaison des réalisations silicium BULK et PD-SOI</b>	<b>120</b>
4.3.1. Nombre d'interrupteurs de puissance unitaires NSW	120
4.3.2. Vitesse et puissance dynamique mesurées du LDPC	126
4.3.3. Puissances statiques mesurées	129
4.3.3.1. Interrupteurs de puissance passants : LDPC alimenté	129
4.3.3.2. Interrupteurs de puissance bloqués : LDPC éteint	130
<b>4.1. Conclusion</b>	<b>132</b>
<b>5. Les bascules de rétention en PD-SOI</b>	<b>135</b>
<b>5.1. Les Bascules de rétention</b>	<b>136</b>
5.1.1. Description	136
5.1.2. Rôle de la bascule de rétention	137
5.1.3. Fonctionnement de la bascule de rétention	138
5.1.4. Environnement de la bascule de rétention	140
5.1.5. La phase de réveil et l'effet transitoire parasite	141
<b>5.2. Suppression de l'effet transitoire</b>	<b>143</b>
<b>5.3. Amélioration pour la technologie PD-SOI</b>	<b>145</b>
5.3.1. Etude de structures pour réduire le courant de fuite	145
5.3.2. Comparaison des différentes implémentations	147
5.3.2.1. Surface	147
5.3.2.2. Vitesse	149
5.3.2.3. Consommation statique	151
<b>5.4. Implémentation optimale</b>	<b>152</b>
➤ <b>Conclusion</b>	<b>155</b>
<b>Brevet et publications</b>	<b>159</b>
<b>Glossaire</b>	<b>161</b>



## ➤ Introduction

L'évolution des technologies décrites par la célèbre loi de Moore [Moo'65] a permis, depuis presque un demi-siècle, d'augmenter sans discontinuité les performances et les fonctionnalités des produits utilisant des circuits intégrés. Un des facteurs essentiels de ce progrès a été l'augmentation de la vitesse de calcul poussée, en particulier, par la réduction des dimensions des composants qui entraîne un courant de saturation plus élevé et des capacités parasites plus faibles. La réduction des géométries apporte également une plus grande densité d'intégration ( $\times 2$  par unité de surface tous les 2 ans), permettant l'accroissement de la complexité fonctionnelle des puces. Cette évolution a structuré la compétition entre les différents acteurs du marché des semi-conducteurs, les produits proposant de plus en plus de fonctionnalités à une vitesse de plus en plus élevée. Toutes les applications ont profité de cette évolution, qu'elles soient fixes ou nomades.

La différence entre les applications fixes et mobiles est le type d'alimentation : sur secteur pour les premières, sur batterie pour les secondes. C'est pour cela qu'il existe plus de contraintes de consommation électrique pour les applications mobiles, alors que celles-ci peuvent passer au second plan pour les applications fixes. Cependant, même pour les applications mobiles, les performances et le nombre de fonctionnalités que propose un produit lui permettent de se démarquer de ses concurrents. Il est donc important pour ces applications d'associer les hautes performances des dernières technologies, afin de répondre aux besoins des consommateurs, à des techniques de réduction de la consommation électrique garantissant une autonomie satisfaisante.

La consommation électrique d'un circuit intégré dépend de son activité, de sa taille, de sa fréquence de fonctionnement et de sa tension d'alimentation. La consommation dynamique est liée aux périodes d'activité, ou de fonctionnement, alors que la consommation statique est associée aux périodes d'inactivité, ou de veille. Tandis que la consommation dynamique est due à la commutation des transistors, la consommation statique est due au maintien électrique des niveaux logiques. La période d'activité d'un circuit dépend également de son domaine applicatif. Les applications fixes ont généralement une période d'activité intense et largement supérieure à leur période d'inactivité. Pour les applications mobiles, c'est souvent le contraire, les temps d'inactivité sont supérieurs aux temps d'activité.

Dans les nœuds technologiques de longueur de grille supérieure à 100nm, la réduction de la consommation électrique totale se faisait en réduisant sa principale composante, la

consommation dynamique. Pour les applications mobiles, la réduction de la consommation électrique impliquait déjà une réduction de la consommation statique, en mode veille. Au fil du temps, pour réduire la consommation dynamique, la tension d'alimentation a été progressivement abaissée. Pour que ce ne soit pas pénalisant en termes de performances, la tension de seuil des transistors a également été réduite, afin de maintenir un rapport tension d'alimentation/tension de seuil à peu près constant. Malheureusement, l'abaissement de la tension de seuil a eu pour effet d'augmenter le courant sous le seuil et donc la consommation statique. Depuis les nœuds technologiques de longueur de grille inférieure à 100nm (technologies dites submicroniques), la consommation statique est devenue une composante aussi importante que la consommation dynamique car elle a augmenté de manière exponentielle au fil des technologies. Cette augmentation est la conséquence, à la fois, de l'abaissement de la tension de seuil et de la réduction de la longueur de grille qui rapproche les zones drain et source et amplifie ainsi l'augmentation du courant sous le seuil. Le nombre de produits mobiles sur le marché ayant explosé, tout comme la consommation statique des circuits des dernières technologies, les concepteurs ont dû développer des techniques de réduction des courants de fuites pour pouvoir proposer des circuits de haute performance mais aussi de faible consommation afin de rendre ces produits nomades plus attrayants.

La technologie Silicium sur Isolant (SOI) partiellement désertée (PD-SOI) apporte un gain en performance par rapport à la technologie CMOS sur substrat massif (BULK), pour le même nœud technologique grâce au comportement de ses transistors à body<sup>1</sup> flottant (FB). Cette technologie est bénéfique aux applications fixes et mobiles. Pour les applications fixes, la technologie PD-SOI permet d'aller plus vite avec une densité d'intégration supérieure. Pour les applications mobiles, la technologie PD-SOI offre la possibilité d'une tension d'alimentation plus faible, permettant d'atteindre les mêmes performances, tout en réduisant la consommation dynamique. L'inconvénient de la technologie PD-SOI vient de sa consommation statique qui est plus élevée qu'en technologie BULK à cause de l'effet de body flottant de ses transistors, qui réduit la tension de seuil. Heureusement, il existe dans cette technologie un transistor à body contacté (BC) qui se comporte électriquement de manière similaire au transistor BULK. C'est pour cela que les techniques de réduction de la consommation statique permettent de mettre en concurrence les technologies PD-SOI et BULK sur tous les aspects, spécialement pour les applications mobiles. La technologie PD-

---

<sup>1</sup> *body* : substrat individuel et isolé de chaque transistor SOI

SOI peut s'appuyer sur des transistors BC pour réduire la consommation statique et sur des transistors FB pour leur rapidité et leur plus faible consommation dynamique.

La technologie PD-SOI est donc une technologie faible consommation et haute performance. Un circuit numérique en technologie PD-SOI peut être issu d'une conception spécifique en PD-SOI ou d'un portage direct à partir du BULK (circuit conçu en BULK puis fabriqué en PD-SOI). Ces portages directs ont montré un rendement fonctionnel équivalent à celui de la technologie sur substrat massif pour une fréquence de fonctionnement plus élevée ou une consommation dynamique plus faible. L'inconvénient de ce portage réside dans le fait que tous les transistors deviennent par construction de type FB, transistor qui fuit plus que son homologue BULK. Cela induit, pour des technologies submicroniques, une consommation statique naturellement plus élevée en PD-SOI qu'en BULK.

Pour adresser ce problème, les travaux de cette thèse ont pour objectif l'optimisation des techniques de réduction de la consommation statique en technologie PD-SOI 65nm basse consommation. Une des techniques de réduction de la consommation statique est la technique MTCMOS basée sur l'insertion d'interrupteurs de puissance. Les interrupteurs de puissance sont des transistors qui permettent de transmettre ou de couper l'alimentation de la partie logique, afin de réduire les courants de fuites pendant les périodes d'inactivité. Cette technique est utilisée en CMOS BULK depuis le nœud 65nm.

Le premier chapitre de ce mémoire introduit la technologie SOI, de la fabrication au comportement électrique du transistor. Les équations spécifiques à cette technologie ainsi que l'effet de body flottant y sont présentés. Cet effet impacte dynamiquement et statiquement la tension de seuil, et par conséquent les courants des transistors. Son impact sur les comportements statiques et dynamiques des transistors est discuté.

Le second chapitre aborde la réduction de la consommation électrique en technologies BULK et PD-SOI. Différentes techniques de l'état de l'art y sont présentées, que ce soit pour réduire la consommation dynamique, la consommation statique ou les deux. Une grande partie de ces techniques est générique et indépendante de la technologie utilisée. D'autres techniques, par contre, sont spécifiques aux technologies : BULK ou PD-SOI. Un bilan est dressé sur ces différentes techniques insistant sur les avantages et inconvénients qu'elles présentent pour la technologie PD-SOI. Il en résulte que la technique la plus bénéfique, dans notre cas en PD-SOI, est la technique MTCMOS.

Une optimisation, pour la technologie 65nm PD-SOI, de la technique MTCMOS basée sur les interrupteurs de puissance est proposée dans le chapitre 3. C'est la technique qui permet de réduire au mieux les courants de fuite tout en conservant les avantages de la



technologie PD-SOI et d'une logique haute performance de type FB. Un nouveau facteur de sélection et de dimensionnement des interrupteurs de puissance est proposé dans le but de sélectionner la solution la plus intéressante en termes de consommation statique, de performances et de surface. A partir de la solution optimale, une solution innovante, appelée Auto-DTMOS, est proposée ; elle est basée sur des transistors de type BC ayant une polarisation auto-adaptative du body. Cette solution a été développée et validée sur silicium pour être finalement comparée à la solution initiale. Elle apporte une réduction de 20% de la résistance équivalente des interrupteurs de puissance, pour le même courant de fuite. Elle nécessite donc moins d'interrupteurs de puissance, pour la même résistance équivalente, ce qui en fait une solution très intéressante pour réduire le courant de fuite. Elle permet aussi d'insérer plus de capacités de découplage sur le rail d'alimentation, ce qui lisse la tension d'alimentation et rend plus robuste le circuit.

Le 4<sup>ème</sup> chapitre présente la validation de cette technique sur un circuit numérique complexe de type LDPC. La description du bloc logique, la méthodologie d'implémentation de la technique MTCMOS ainsi que le choix du nombre d'interrupteurs de puissance y sont décrits. Les résultats de mesure permettent de déterminer le gain apporté par la technologie PD-SOI, ainsi que le gain issu de la conception par rapport au BULK en termes de consommation statique. Le gain technologique apporte une fréquence 20% plus rapide en PD-SOI par rapport au BULK pour la même tension d'alimentation, et une réduction de la consommation dynamique de 30%, pour la même fréquence. Le gain lié à la conception en PD-SOI, en introduisant la nouvelle solution, Auto-DTMOS, permet de diviser par 2 la consommation statique en mode veille, pour les mêmes performances atteintes dans le mode actif.

Le dernier chapitre évoque l'optimisation en PD-SOI de la bascule de rétention, bloc indispensable à associer à la technique MTCMOS. Cette bascule de rétention permet de conserver les états logiques importants, même lorsque les interrupteurs de puissances sont éteints. Elle a été étudiée et conçue pour être la plus robuste possible (un effet transitoire indésirable y est supprimé). La bascule est optimisée avec pour objectifs la réduction de la consommation statique et l'amélioration des performances.

En conclusion, les principaux apports de ce travail sont résumés, en rappelant les gains et résultats obtenus. Les pistes d'amélioration sont tracées et une présentation des perspectives à venir sur la technologie Fully Depleted SOI en 28nm est finalement livrée.

[Moo'65] Gordon E. Moore, "Cramming More Components Onto Integrated Circuits," *Electronics*, vol. 38, April 1965





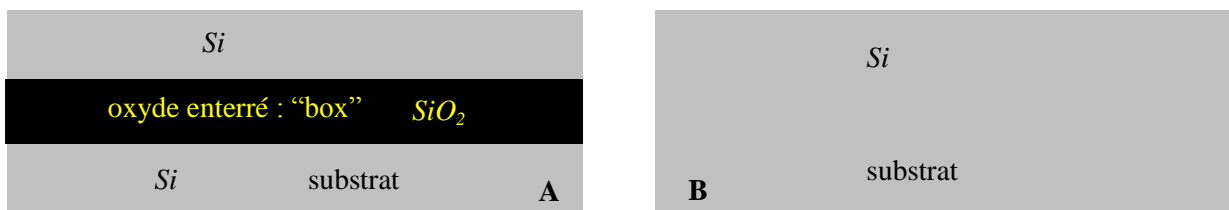
# 1. Les technologies SOI

*Les technologies Silicium-Sur-Isolant « SSI » ou « SOI »(en anglais pour Silicon-On-Insulator) sont des technologies développées pour des applications faible consommation « Low Power » ou haute performance « High Performance ». Ces technologies, et plus particulièrement la technologie partiellement désertée, « PD-SOI », sont présentées dans ce chapitre, de la fabrication des tranches de silicium, en passant par le comportement électrique du transistor unitaire jusqu'à la comparaison de circuits avec la technologie silicium sur substrat massif dite « BULK ».*

## 1.1. Substrat et Transistor

### 1.1.1. Substrat

D'un point de vue physique et mécanique, un substrat SOI [Col'04] est constitué de deux couches de silicium, « Si », isolées l'une de l'autre par une couche d'oxyde enterrée (« box » en anglais pour buried oxide), contrairement au substrat massif classique, ou « BULK », qui n'est constitué que de silicium. Ces substrats, SOI et BULK, sont représentés respectivement figure .a et figure .b.



**figure : Vue du substrat des technologies SOI (A) et BULK (B)**

Le principal et plus utilisé des procédés de fabrication industrielle de substrats SOI reste aujourd'hui le découpage intelligent ou « smart cut » [Del'09]. L'intérêt est d'obtenir une couche de silicium supérieure mince et isolée.

### 1.1.2. Transistor

La figure représente, sur SOI et sur technologie BULK, le Transistor à Effet de Champ, Métal Oxyde Semi-conducteur, « MOSFET » pour Metal Oxide Semiconductor Field Effect Transistor.

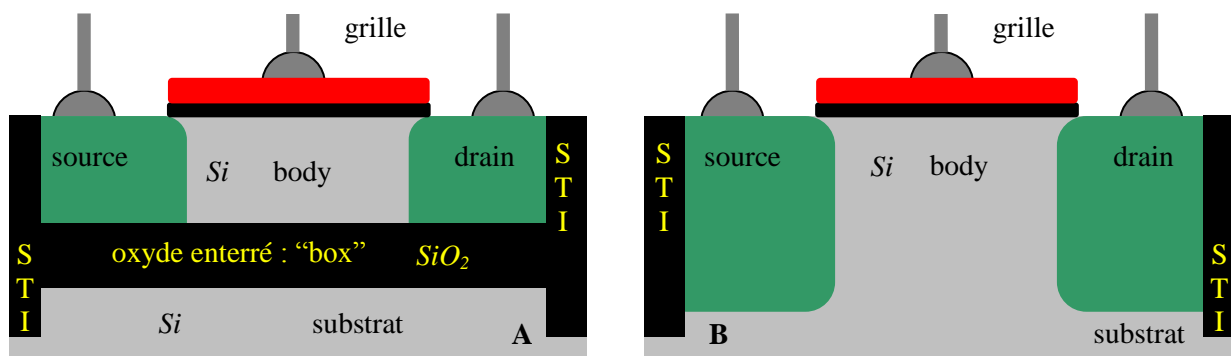


figure : Vue d'un transistor MOSFET en technologies SOI (A) et BULK (B)

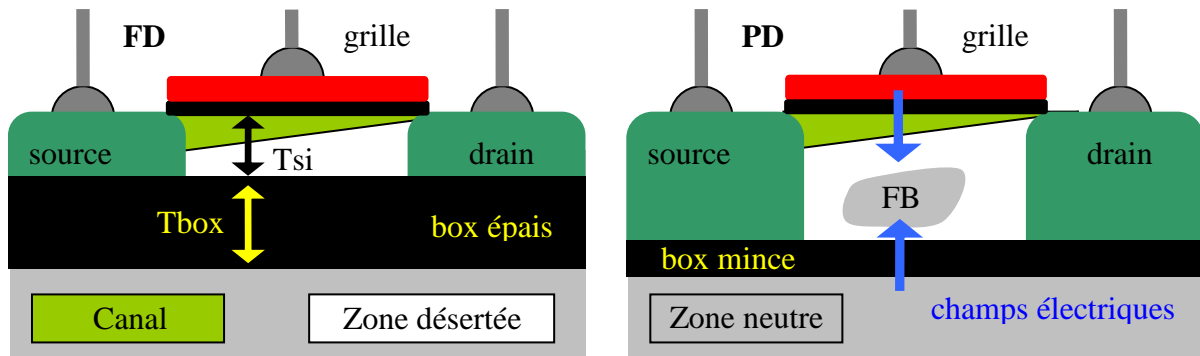
La face avant, étant la couche de silicium la plus mince, est destinée à accueillir les composants du circuit. L'oxyde enterré permet lui, non seulement d'isoler ces composants de la face arrière (protection des radiations et isolation électrique), mais également entre eux, grâce à l'utilisation de l'isolation latérale « STI » (shallow trench isolation), supprimant ainsi l'effet latch-up [Sch'03]. La spécificité de la technologie SOI réside dans le fait que la zone entre le drain et la source appelée « body » est isolée, car non-accessible, ce qui n'est pas le cas en BULK.

### 1.1.3. Pourquoi le SOI ?

Initialement conçue dans une optique d'utilisation militaire et spatiale pour ses aspects de robustesse notamment grâce à son immunité face aux radiations [Pel'03], l'objectif aujourd'hui principal de cette technologie est d'apporter une solution en termes de haute performance et de faible consommation [Ber'00]. Ceci est obtenu grâce à l'isolation du body et grâce à la réduction des capacités de jonction, qui permettent d'augmenter les performances et de diminuer la consommation dynamique. De plus, la prise substrat n'est plus nécessaire, [Pel'05], ce qui permet une densité d'intégration légèrement meilleure.

### 1.1.4. Transistors PD-SOI et FD-SOI

Il existe plusieurs types de technologies SOI, et les transistors de chaque technologie n'ont pas les mêmes caractéristiques électriques:



**figure : Transistor SOI entièrement et partiellement désertée (FD & PD).**

Comme le montre la figure , en fonction du dopage du canal et selon l'épaisseur du silicium « Tsi », le transistor d'une technologie SOI peut être défini comme entièrement ou partiellement déserté (respectivement « FD » ou « PD » pour fully ou partially depleted). L'épaisseur du film de silicium agit directement sur les caractéristiques électriques [Raj'95]. Il est alors question de film de silicium mince pour FD ou épais pour PD.

La différence de potentiel entre la grille et la source,  $V_{gs}$ , agit en profondeur dans la zone sous la grille, entre le drain et la source. A faible  $V_{gs}$ , inférieure à  $V_{th}$ , c'est une zone désertée ou de faible inversion qui va se créer. A fort  $V_{gs}$ , supérieure à  $V_{th}$ , une zone de forte inversion ou « canal » va progressivement remplacer la zone de désertion. Ce phénomène est lié au champ électrique induit par  $V_{gs}$ , champ électrique qui s'atténue avec la distance aux bornes d'une structure MOS. Au-delà d'une certaine profondeur, lorsque le champ électrique devient nul, il existe une zone qui conserve sa neutralité électrostatique. Cela définit, dans l'axe de la profondeur (de la grille au body) et pour un fort  $V_{gs}$ , une zone de forte inversion (canal), une zone de désertion et une zone neutre [Sin'98].

Dans le cas des technologies entièrement désertées FD, il n'existe pas de zone neutre, l'épaisseur  $T_{si}$  est inférieure à la profondeur de désertion. En revanche pour les technologies partiellement désertées PD, l'épaisseur du silicium étant supérieure à la zone de désertion, une zone neutre apparaît, pouvant se transformer en zone d'accumulation de charges. L'accumulation de charges dans la zone neutre fait augmenter son potentiel, modifiant ainsi les caractéristiques électriques du MOSFET (tension de seuil) : c'est l'effet du body flottant ou « FBE » pour Floating Body Effect en anglais. Ce body flottant « FB » se situe au-delà du champ électrique contrôlé par  $V_{gs}$ . Il est proche de la zone en contact avec l'oxyde enterré.

La zone de désertion maximale  $x_{dmax}$  est donnée par l'équation . Dans le cas de la technologie SOI Partiellement Désertée, PD-SOI : l'épaisseur d'oxyde  $T_{si} > 2 * x_{dmax}$ . Dans le cas de la technologie totalement désertée FD-SOI :  $T_{si} < x_{dmax}$  ;  $\Phi_F$  est le potentiel de Fermi,

$N_a$  est la concentration de dopage dans la région P (cas du NMOS),  $\epsilon_{si}$  est la permittivité du silicium et  $q$  est la charge d'un électron.

$$x_{d\max} = \sqrt{\frac{4\epsilon_{si}\Phi_F}{qN_a}}$$

**équation**

L'épaisseur de l'oxyde enterré « Tbox » joue aussi un rôle déterminant. En effet, le contrôle du body par le champ électrique de la face arrière peut entraîner une modulation des charges dans celui-ci. Ce champ électrique est fonction de l'épaisseur de l'oxyde enterré [Ban'97] et du potentiel appliqué sur la face arrière [Col'85]. L'oxyde enterré peut être mince ou épais [Kim'09] et selon son épaisseur, les capacités observées du point de vue de la grille et du body vont également évoluer mais de manière peu significative. STMicroelectronics propose une technologie PD-SOI 65nm de faible consommation (Low-Power). Pour cette technologie, le film de silicium est de 70nm, le substrat de la face arrière est polarisé à GND et l'oxyde enterré est suffisamment épais, 140nm, pour que le champ électrique de la face arrière soit quasi nul et n'ait aucun impact sur le potentiel de body.

L'avantage de la technologie SOI est sa compatibilité avec la plateforme de conception BULK [Fla'10]. La technologie PD-SOI reste la mieux adaptée pour les technologies supérieures au nœud technologique 32nm. La technologie FD-SOI se présente logiquement comme son héritière. Il reste néanmoins de belles années à la technologie PD-SOI, car la technologie FD-SOI n'est pas encore arrivée au stade de production.

## 1.2. Comportement électrique des transistors PD-SOI

La technologie SOI de type entièrement désertée, « FD-SOI » présente des caractéristiques électriques intéressantes, avec une excellente pente sous le seuil et un meilleur contrôle du canal par la grille qui n'engendre pas d'effet de body flottant [Sak'06] contrairement à la technologie SOI de type partiellement désertée, « PD-SOI ». La technologie PD-SOI, quant à elle, offre un gain en performance lié à l'effet de body flottant.

D'un point de vue industriel, la technologie FD-SOI est envisagée pour les nœuds très avancés, pour lesquels ses avantages deviennent déterminants.



### 1.2.1. Réduction des capacités de jonction

Ce qui est commun aux technologies PD-SOI ou FD-SOI, oxyde enterré mince ou épais, c'est le gain en consommation dynamique lié à la réduction des capacités de drain et de source. La couche isolante a permis de réduire en profondeur la surface des diffusions drain et source entraînant une réduction des capacités de jonction [Kri'98]. Il est dès lors possible d'augmenter la fréquence, en comparaison avec le BULK, pour une même consommation dynamique [Col'94].

Le tableau représente les équations de la capacité de jonction par unité de surface des technologies BULK et SOI (respectivement l'équation et l'équation ). Le ratio  $C_{BULK}/C_{SOI}$ , en typique, peut varier entre 4 et 7 [Col'04].  $V_D$  est la différence de potentiel aux bornes de la capacité,  $\Phi_o$  est le potentiel de jonction,  $N_d$  et  $N_a$  sont les concentrations de dopage dans les régions N et P,  $\epsilon_{si}$  est la permittivité du silicium,  $q$  est la charge d'un électron et  $C_{BOX}$  la capacité de l'oxyde enterré par unité de surface.

Capacité de jonction par unité de surface	(F/cm <sup>2</sup> )
BULK	$C_{BULK} = \sqrt{\frac{q\epsilon_{si}N_aN_d}{2(N_a + N_d)}} \frac{1}{\sqrt{\Phi_o - V_D}}$ <p>équation</p>
SOI	$C_{SOI} = \frac{C_{BOX}}{\sqrt{1 + \frac{2C_{BOX}^2 V_D}{qN_a\epsilon_{si}}}}$ <p>équation</p>

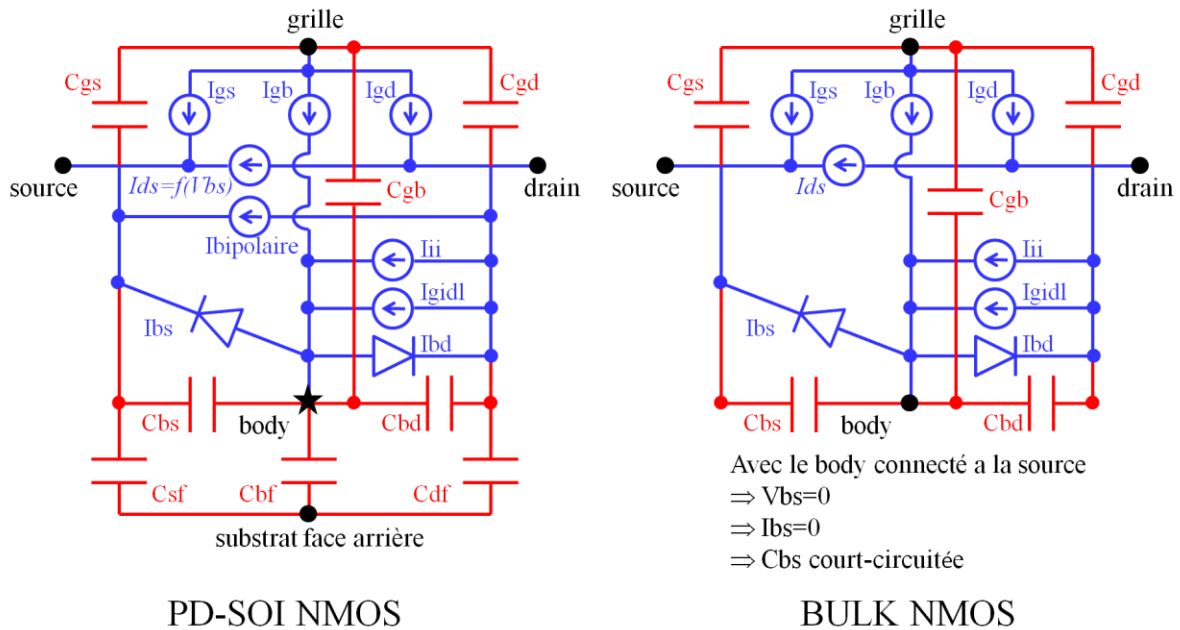
tableau : Capacité de jonction en BULK et SOI

[Wei'98] présente l'impact de la réduction des capacités sur les caractéristiques et les performances des circuits CMOS en technologie PD-SOI.

### 1.2.2. L'effet de body flottant

Le body flottant des transistors de la technologie PD-SOI modifie le comportement électrique de ceux-ci par rapport aux transistors de la technologie BULK. La figure montre

un schéma électrique interne, équivalent et simplifié au premier ordre, du NMOS de type PD-SOI et BULK. Les capacités de jonction sont réduites en PD-SOI, laissant place à d'autres capacités liées à l'oxyde enterré, beaucoup plus faibles et peu influentes sur le comportement dynamique. Le modèle électrique équivalent n'est pas le même en PD-SOI et en BULK. En SOI, il est impacté par un body non plus connecté à un potentiel fixe, comme en BULK, mais flottant.



**figure : Schéma électrique simplifié du NMOS PD-SOI et BULK**

L'isolation du body entraîne la formation d'une zone flottante en profondeur en PD-SOI où accumulation de charges, phénomènes de couplages capacitifs et effet d'histoire, vont se produire. Le body flottant est également la cause d'un effet appelé « kink », qui amplifie le courant  $I_{ds}$  en fonction de  $V_{bs}$  et de  $V_{ds}$ , ainsi qu'un effet bipolaire. Tous ces phénomènes sont présentés dans cette section mais leurs impacts électriques seront détaillés dans la partie consacrée à l'étude des courants du MOSFET.

### 1.2.2.1. L'accumulation de charges dans le body flottant

L'accumulation de charges dans le body flottant, illustrée figure , présente l'avantage de réduire la tension de seuil  $V_{th}$  et donc d'augmenter le courant de canal, en modifiant le potentiel du body.

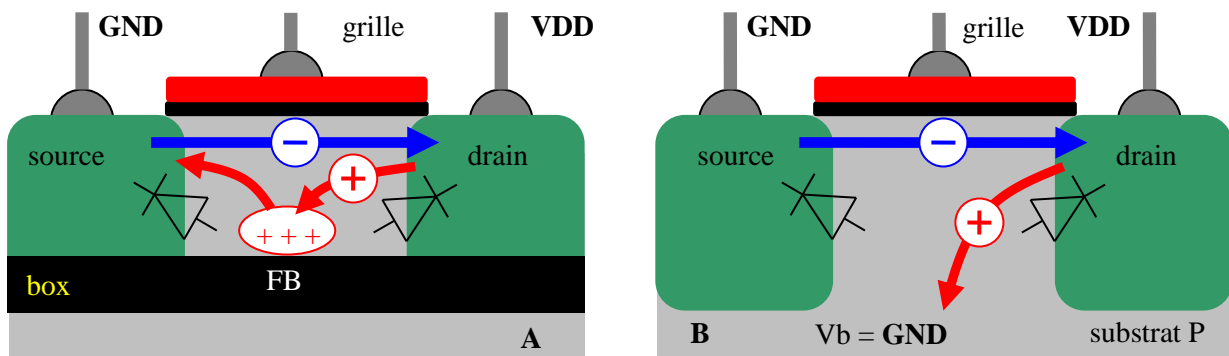


figure : Principe du déplacement des charges dans un NMOS de technologies SOI (A) et BULK (B)

Dans le cas d'un MOS de type N « NMOS », le déplacement des porteurs négatifs du canal se fait de la source au drain. Les trous qui sont proches du drain vont majoritairement se déplacer vers la source mais également vers le body. En BULK, la connexion du body permet une évacuation totale des trous qui vont se recombiner dans le substrat. En PD-SOI, cette évacuation par le body n'étant pas possible, les charges vont s'accumuler dans celui-ci. Cette accumulation augmente la tension de body et donc abaisse la tension de seuil, ce qui a pour effet d'augmenter le courant de canal.

L'augmentation du courant de canal accroît l'accumulation de charges positives dans le body flottant, ce qui continue à élever la tension de body. L'accumulation s'arrête grâce à la diode de jonction entre le body et la source qui devient passante et qui permet l'évacuation des trous du body vers la source. Réciproquement ce phénomène se produit avec des charges négatives pour les transistors de type P « PMOS ».

### 1.2.2.2. Le couplage capacitif

Les phénomènes de couplage capacitif entraînent la modulation des charges dans le body flottant en fonction de la fluctuation des potentiels du drain, de la source et de la grille. Il est clair que le comportement capacitif d'un point de vue du canal ou du body est différent d'une technologie à l'autre. Il est présenté figure .

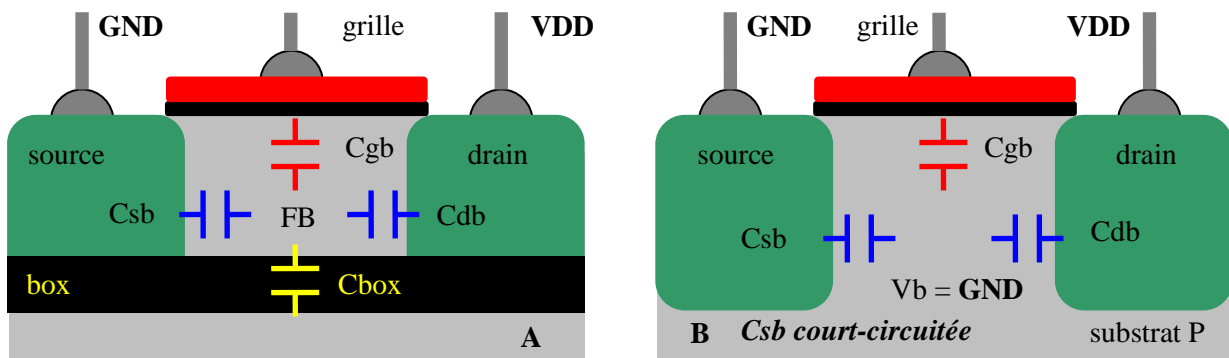
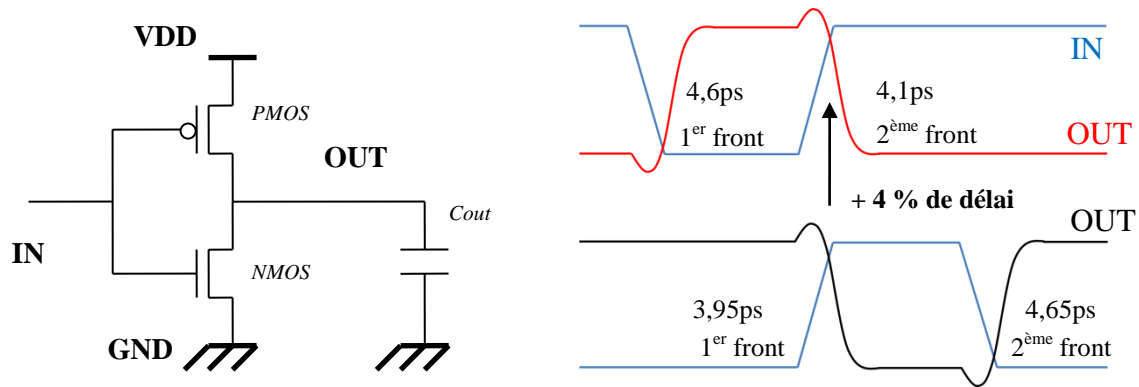


figure : Effets capacitifs dans un NMOS, en technologies SOI (A) et BULK (B)

Pour un NMOS et dans le cas du BULK, le potentiel de body, «  $V_b$  », est fixé à GND. Dans le cas du PD-SOI, le potentiel de body, varie statiquement en fonction des charges accumulées et dynamiquement en fonction des potentiels des nœuds du MOS. Quand  $V_g \uparrow$  alors  $V_b \uparrow$  mais lorsque  $V_d \downarrow$  alors  $V_b \downarrow$ . Si le potentiel  $V_b$  évolue en fonction des différents potentiels du transistor, alors la tension de seuil ainsi que le comportement du transistor évoluent également.

#### 1.2.2.3. L'effet d'histoire

L'accumulation et l'évacuation des charges dans le body peuvent être rapides (ps) ou lentes (ms) et engendrer une variation du potentiel du body induisant également une variation de la tension de seuil  $V_{th}$  en fonction du temps. Ces variations sont liées aux différents états de polarisation antérieurs. En fonction des polarisations statiques et du comportement dynamique du transistor, la tension de seuil instantanée et moyenne de celui-ci va varier. Ce phénomène s'appelle l'effet d'histoire [Nar'01]. Ainsi un transistor n'aura pas les mêmes caractéristiques électriques selon son état initial ou selon son rapport cyclique et sa fréquence [Lio'03] puisque sa tension de seuil ne sera pas la même. La figure présente un exemple de variation de délai que peut entraîner l'effet d'histoire. Lors de la conception d'un circuit, cette variation n'est pas à négliger.



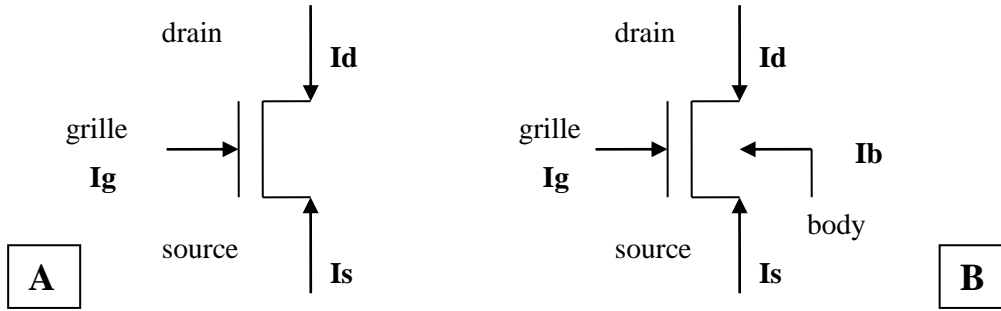
**figure : Impact de l'effet d'histoire sur le délai d'un inverseur sur front montant de 5ps à  $V_{DD}=1,2V$  à  $25^{\circ}C$ .**

En considérant deux commutations sur le même front mais avec des états initiaux différents, on peut observer jusqu'à 4% de dégradation de délai pour un inverseur simple en 65nm PD-SOI Haut-Vt « HVt ». Ceci s'explique par un potentiel de body,  $V_b$ , différent dans les deux cas et qui est lié à l'accumulation lente et rapide des charges. Dans un cas, le NMOS est initialement bloqué avec  $V_{ds}=V_{DD}$  et  $V_b$  proche de  $V_{DD}/2$ , et dans l'autre, le NMOS est initialement passant en mode linéaire avec un  $V_{ds}$  quasi nul et  $V_b$  valant à peine une centaine de millivolts. La différence des potentiels de body,  $V_b$ , dans les deux cas induit une différence des tensions de seuil et donc des performances.

Les principaux phénomènes liés au body flottant ont été décrits dans cette première partie de chapitre. Dans la seconde partie, l'impact de ces phénomènes sur les courants actifs et de fuite du transistor est présenté.

### 1.2.3. Les courants du MOSFET

Il existe 4 potentiels accessibles pour le transistor MOSFET de technologie BULK, alors qu'il n'en existe que 3 pour son homologue de technologie SOI, comme le montre la figure .  $I_d$ ,  $I_g$ ,  $I_s$ ,  $I_b$  sont respectivement les courants de drain, grille, source et body (exclusivement pour le BULK).



**figure : Courants du MOSFET de technologie PD-SOI (A) et BULK (B)**

#### 1.2.3.1. Courant actif

Lorsque le transistor est actif, en mode linéaire ou saturé, le courant dominant est le courant de canal  $I_{ds}$ , les autres courants étant fortement négligeables. C'est ce courant qui permet la commutation des portes logiques : plus il est élevé plus la commutation est rapide. La valeur de ce courant en forte inversion dépend de l'équation [Mis'08].

Le courant actif  $I_{ds}$ , défini par l'équation , circule à travers le canal de forte inversion. Les paramètres sont  $\mu_n$  la mobilité,  $C_{ox}$  la capacité de l'oxyde de grille,  $W$  et  $L$  respectivement la largeur et longueur du MOSFET,  $V_{th}$  la tension de seuil et  $V_{gs}$  et  $V_{ds}$  les différences de potentiel entre la grille et la source ainsi qu'entre le drain et la source.

$$I_{ds} = \frac{W}{L} * \mu_n C_{ox} * \left( V_{gs} - V_{th} - \frac{V_{ds}}{2} \right) * V_{ds}$$

**équation**

En PD-SOI, les phénomènes précédemment présentés comme le couplage capacitif [Pel'02] par la grille (figure ), ainsi que l'accumulation de charges du nœud flottant due au courant d'ionisation par impact Iii [Su'01] (figure ), vont abaisser dynamiquement et statiquement la tension de seuil  $V_{th}$  et donc permettre d'augmenter favorablement le courant actif  $I_{ds}$  en comparaison avec le BULK. C'est l'avantage de l'effet de body flottant du MOSFET PD-SOI. Comme la tension de seuil est liée à la racine carrée de la différence de potentiel  $V_{bs}$  par la relation suivante (équation ), et que l'effet de body flottant amène  $V_{bs}$  à augmenter, le courant actif du MOS PD-SOI est plus élevé qu'en BULK.

$$V_{th} = V_{th0} - \beta \times \sqrt{V_{bs}}$$

équation

$V_{th0}$  est la tension de seuil de référence lorsqu'aucune charge n'est accumulée dans le body ( $V_{bs}=0$ ) et  $\beta$  est le coefficient d'effet de body.

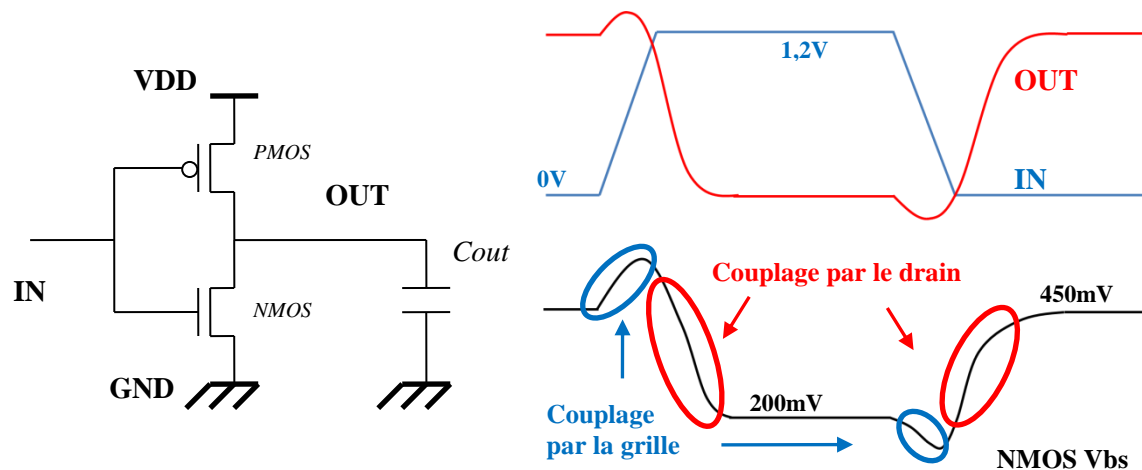


figure : Couplage capacitif du body du NMOS d'un inverseur à  $V_{DD}=1,2V$  à  $25^{\circ}C$

Le couplage capacitif par la grille permet de dynamiquement augmenter le potentiel du nœud flottant  $V_b$  et donc de réduire  $V_{th}$  (dans le cas d'un NMOS), au début de la transition. Le couplage capacitif par le drain ramène le potentiel de body à sa valeur statique. Il est présent en milieu et fin de transition, comme montré dans la figure ci-dessus. Le couplage capacitif permet notamment d'augmenter le courant actif en début de transition.

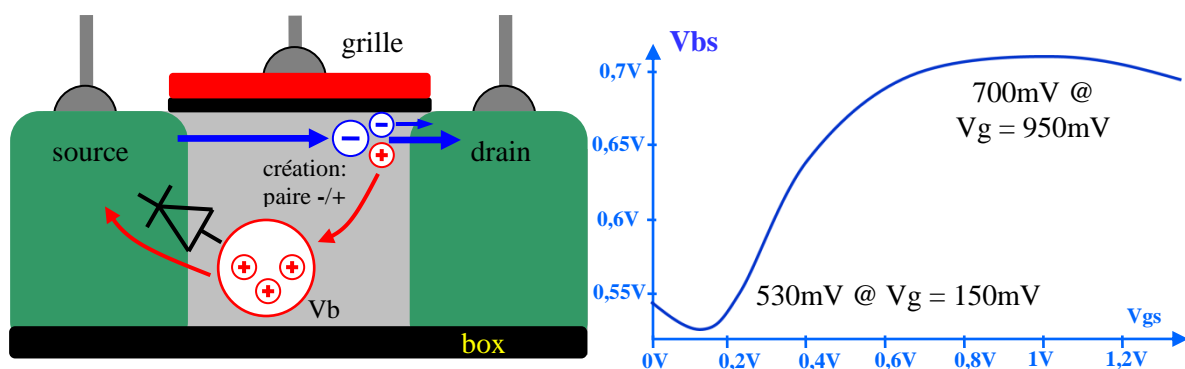


figure : Accumulation de charges dans le body ( $V_{bs}$ ) par courant d'ionisation par impact en fonction du potentiel de grille ( $V_{gs}$ ) pour  $V_{ds}=1,35V$  à  $25^{\circ}C$

Lorsque le transistor est passant, le courant de canal et le champ électrique dans la zone sous la grille entre le drain et le body entraînent la création de paires électrons/trous ;

c'est la formation du courant d'ionisation par impact (Iii). Les trous vont se déplacer dans le body flottant (cas du NMOS) et vont charger celui-ci, ce qui aura pour effet d'augmenter  $V_b$  et donc de réduire  $V_{th}$ . En réduisant  $V_{th}$ ,  $I_{ds}$  augmente ainsi que le courant d'ionisation par impact. Heureusement, l'évacuation des charges du body vers la source via la diode qui existe entre eux et qui devient directe permet de stabiliser ce phénomène. Cet effet d'avalanche est appelé l'effet « kink », il met en évidence une augmentation soudaine du courant de canal en fonction de la tension  $V_{ds}$ .

Il existe également un effet bipolaire parasite [Lin'99] qui apparaît pour des tensions  $V_{gs}$  et  $V_{ds}$  plus élevées. Dans ce cas, le body devient comparable à la base d'un transistor bipolaire avec Iii et le courant de « GIDL » comme courant de base  $I_b$  et la source et le drain respectivement l'émetteur et le collecteur pour un NMOS. En technologie LP (Low Power : faible puissance) 65nm, les tensions d'alimentation typiques et maximales, respectivement de 1,2V et 1,35V, sont bien inférieures aux tensions observées lorsque le phénomène bipolaire se déclenche.

L'effet de body flottant et la réduction des capacités de jonction sont à l'origine du gain en performance et en consommation dynamique de la technologie PD-SOI par rapport à son homologue BULK.

Cet effet flottant a également des conséquences sur la valeur des courants de fuite du transistor en mode inactif, et d'une manière générale sur la consommation statique.

#### 1.2.3.2. Courants de fuite

Concernant les fuites, elles sont devenues de plus en plus importantes, quelle que soit la technologie, en raison de deux phénomènes. Premièrement, afin de réduire la consommation dynamique des circuits intégrés, la tension d'alimentation  $V_{DD}$  a été réduite. La tension de seuil,  $V_{th}$ , a également été réduite [Kea'07] afin de maintenir un ratio  $V_{DD}/V_{th}$  constant, synonyme de maintien des performances. Le second phénomène est la réduction des dimensions ; les zones drain source se sont rapprochées entraînant des effets de canaux courts [Suz'00], représentés figure . Ce rapprochement entre le drain et la source réduit la résistivité du canal sous le seuil (MOS bloqué), ce qui a tendance à augmenter le courant  $I_{ds}$ .



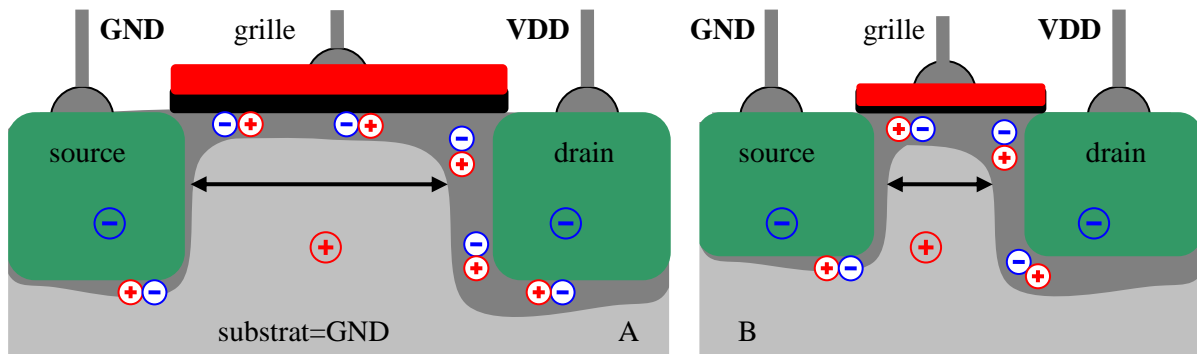


figure : MOSFET a canal long (A) et court (B)

Les courants de fuite sont des courants apparaissant aux bornes du transistor lorsque celui n'est pas en commutation. Dans le cas du NMOS ; lorsqu'il est bloqué avec  $V_{gs}=0$  et  $V_{ds}=VDD$  ou passant avec  $V_{gs}=VDD$  et  $V_{ds}$  quasi nul (de l'ordre de grandeur du micro/milli volt).

Il existe plusieurs courants de fuite dans le transistor BULK et SOI : La figure ainsi que le tableau permettent de résumer et distinguer les courants communs et dépendants des technologies BULK [Roy'03] et PD-SOI [Che'08] à canaux courts.

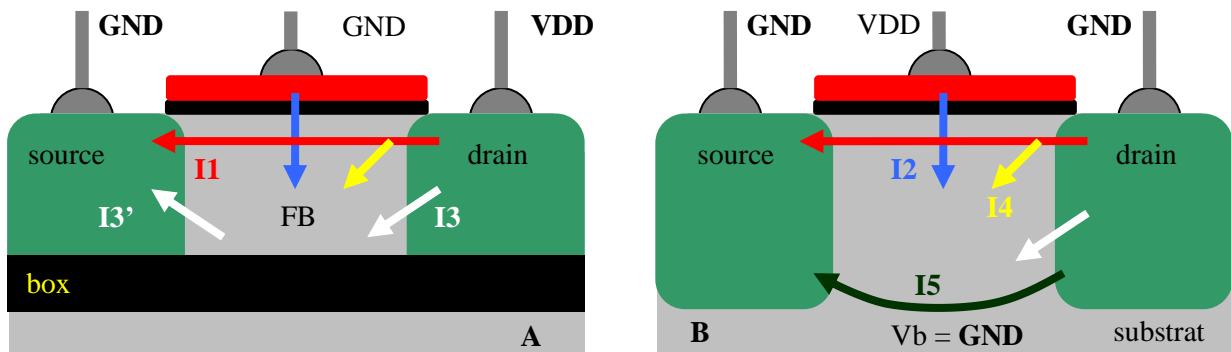


figure : Principaux courants internes d'un MOS de technologies PD-SOI (A) et BULK (B)

Courant	Composante	Commentaire
I1	$I_{DS}$	Courant de canal faible inversion
	$I_{DIBL}$	« Drain induced barrier lowering »
I2	$I_{TUN}$	Courant tunnel à travers l'oxyde mince de grille
	$I_{INJ}$	Courant dû à l'injection de porteurs chauds
I3	$I_{DIODE}$	Courant indirect de la diode de jonction
I3'	$I_{DB}$	Courant direct de la diode de jonction
I4	$I_{GIDL}$	« Gate induced drain leakage »
I5	$I_{PT}$	Courant de perçage (BULK)

**tableau : Courants internes détaillés d'un MOS de technologies PD-SOI et BULK.**

Le courant sous le seuil  $I_{DS}$  est le courant de faible inversion sous la grille [Ada'01] lorsque  $V_{gs}=0$ . Il est majoritairement responsable des fuites et donc de la consommation statique. Plusieurs modèles analytiques ou physiques ont été proposés pour ce courant et pour des technologies submicroniques [Aga'05] ; tous démontrent une dépendance exponentielle de  $I_{ds}$  par rapport à l'inverse de  $V_{th}$ . Laquelle réduction de  $V_{th}$  a dramatiquement fait augmenter le courant  $I_{ds}$  et donc la consommation statique.

Le courant « Drain Induced Barrier Lowering » (DIBL) [Guo'06] est le courant lié à la réduction de la barrière induite par le potentiel de drain. Ce phénomène impacte la tension de seuil  $V_{th}$ . A faible  $V_{ds}$ , le DIBL est nul, dans ce cas  $|V_{th}|$  est égale à  $V_{tlin}$ . A l'opposé à fort  $V_{ds}$ , le DIBL est à son maximum,  $|V_{th}|$  vaut  $V_{tsat}$ . Le DIBL réduit la tension de seuil  $V_{th}$ ,  $V_{tsat}$  étant inférieure à  $V_{tlin}$ .

Le courant « Gate Induced Drain Leakage » (GIDL) [Ade'07] est un courant à effet tunnel créé dans la zone de chevauchement entre la grille et le drain et lié à la tension  $V_{gs}$ .

Le courant de perçage est un courant de fuite faible et très limité en SOI. Il est lié au rapprochement en profondeur des zones drain et source [Kot'02].

Le courant tunnel de grille à travers l'oxyde ainsi que le courant d'injection de porteurs chauds, plus faibles en SOI qu'en BULK, restent des courants de fuites minoritaires en 65nm [Yeh'02].

La figure met en évidence les courants en fonction de la tension de grille du transistor NMOS à tension de seuil élevée, « HVt ».

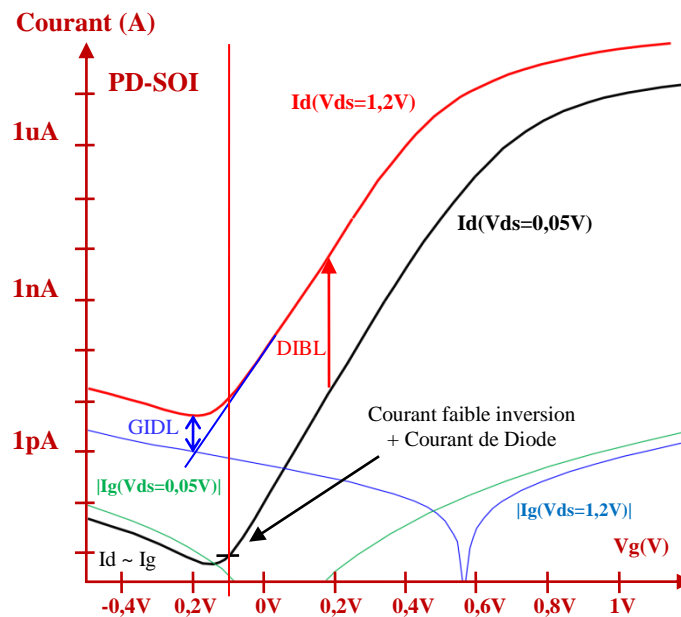


figure : Courant total  $I_d$  &  $|I_g| = f(V_g)$  pour  $V_{ds}=0,05V$  &  $1,2V$  d'un NMOS HVt  $W=0,2\mu m$   $L=60nm$  à  $25^\circ C$

La figure montre une comparaison entre les deux technologies BULK et SOI en 65nm. A noter qu'à faible  $V_{ds}$ , l'accumulation de charges dans le body devient de plus en plus faible et l'effet body flottant disparaît. En effet la tension de body  $V_b$  est toujours proche en statique des potentiels  $V_d$  et  $V_s$ .

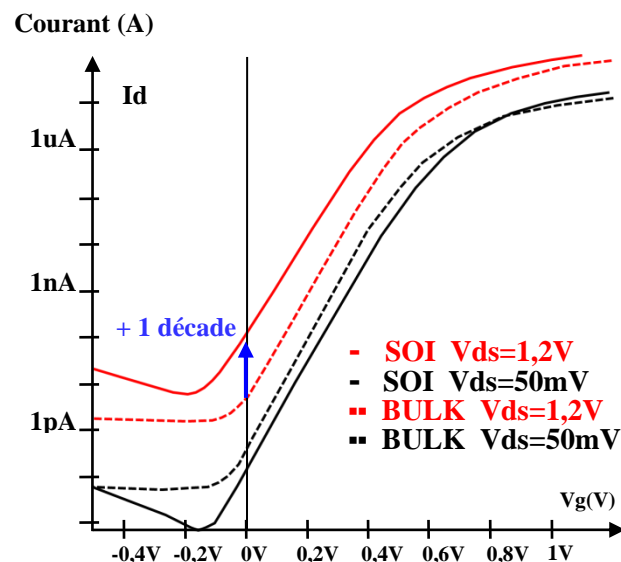


figure : Courant  $I_d$  pour  $V_{ds}=0,05V$  &  $1,2V$  d'un NMOS HVt  $W=0,2\mu m$   $L=60nm$  à  $25^\circ C$

En PD-SOI l'accumulation de charges dans le body flottant amène une réduction de la tension de seuil  $V_{th}$ , de quelques pourcents, ce qui augmente le courant sous le seuil [Cai'07]

et donc les courants des fuites. La consommation statique d'un circuit PD-SOI peut être 5 à 10 fois supérieure à celle du même circuit en BULK. C'est aujourd'hui l'un des problèmes majeurs de la technologie PD-SOI [Lho'06] pour des applications basse consommation.

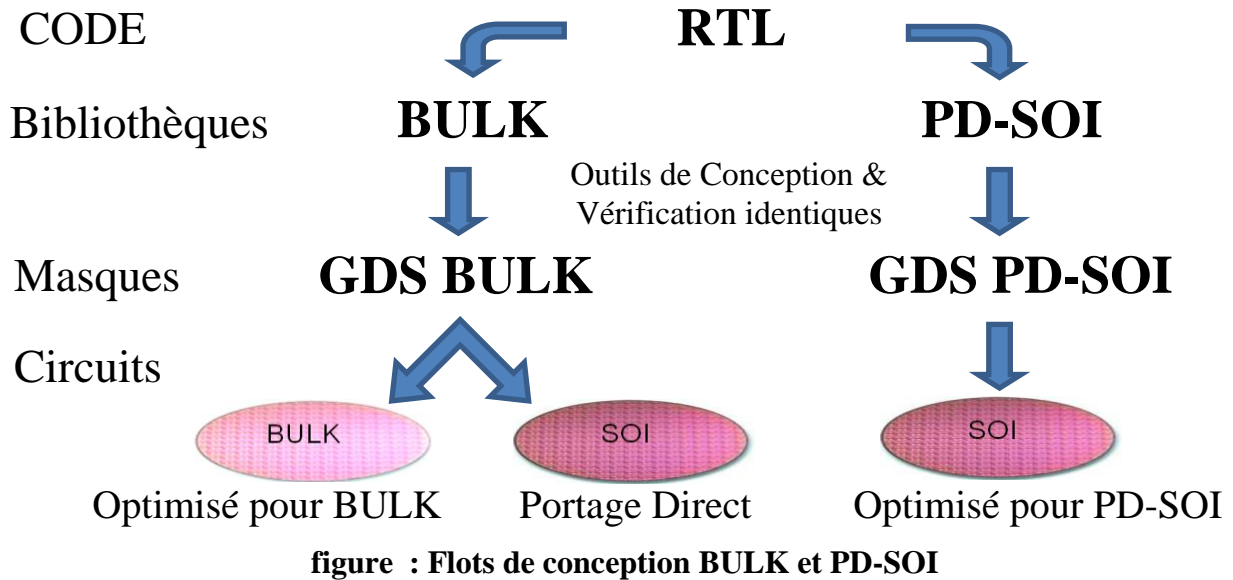
### 1.2.4. Comportement en température

La température ainsi que l'auto-échauffement (phénomène lié à l'isolement diélectrique du transistor empêchant le transfert de chaleur [Su'01]) ont une influence importante sur les performances et la consommation dynamique et statique du transistor. En PD-SOI, l'effet du body flottant se retrouve réduit par la température. La diminution de l'effet flottant, impactant sur la tension de seuil, permet de réduire le ratio entre les courants de fuite PD-SOI et ceux du BULK pour des températures élevées [Fos'98]. Car en BULK, les jonctions drain, source avec le body sont plus grandes : ce qui signifie qu'elles sont plus fuyantes avec l'augmentation de la température qu'en PD-SOI.

### 1.2.5. Comparaison entre les technologies BULK et PD-SOI

Comme cela a été précisé en début de ce chapitre, la technologie PD-SOI est compatible avec la technologie BULK en termes de flot de conception. Les études et réalisations ont montré qu'un portage direct ou dit « à l'aveugle » du BULK vers PD-SOI apporte systématiquement un gain en consommation dynamique et en vitesse et ceci pour le même rendement fonctionnel [Mat'01].

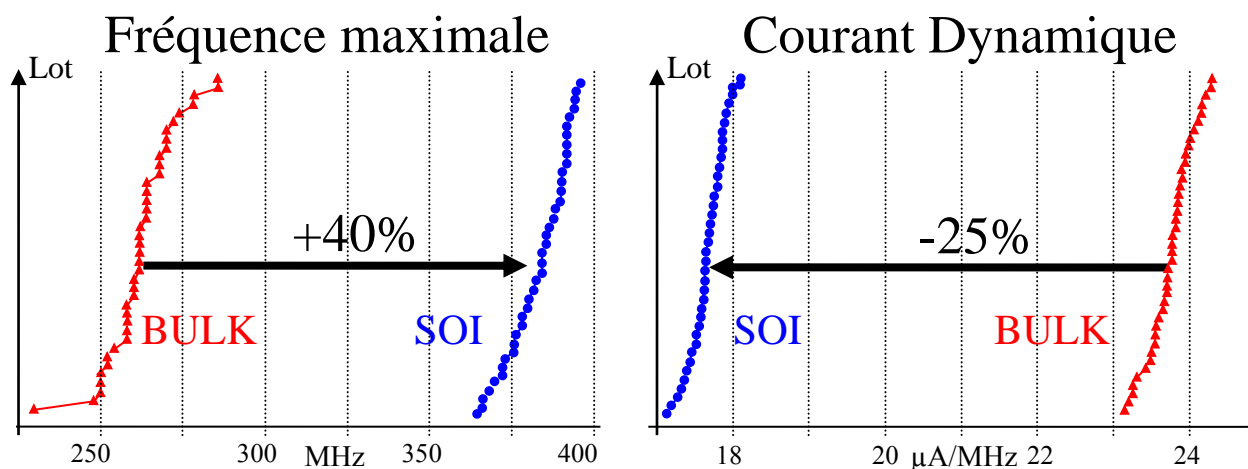
Cela signifie qu'un circuit numérique conçu pour une technologie BULK avec des outils dédiés à cette technologie (tels que la plateforme de conception, les outils de vérifications et de simulations) peut néanmoins être fabriqué sur une technologie PD-SOI. Quelques étapes intermédiaires permettent, sans aucune modification du jeu de masques nécessaires à la fabrication du circuit, de vérifier que les contraintes de délai sont toujours respectées. La figure met en évidence les différents flots de conception possibles pour la fabrication d'un circuit PD-SOI. Certains blocs conçus en BULK ne sont pas optimisés pour la technologie PD-SOI. Par exemple, le portage direct des plots d'entrées/sorties ne permettra pas de respecter les mêmes spécifications en termes de protection contre les décharges électrostatiques.



L'inconvénient du portage direct réside aussi dans le fait qu'il ne permet pas une optimisation de circuit en PD-SOI. Ainsi un tel portage en technologie 0,18 $\mu$ m apporte un gain en performance sur un circuit « ALU » (unité arithmétique logique) de 16% contre 21% lorsqu'il est conçu et optimisé en PD-SOI [Mat'01].

#### 1.2.5.1. Portage direct d'une mémoire SRAM

La figure montre les gains en vitesse et en consommation dynamique d'une mémoire SRAM de technologie 65nm.



**figure : 65nm SRAM : Comparaison Fréquence Maximale et Courant Dynamique entre BULK et PD-SOI à VDD=1,2V et 25°C.**

La technologie PD-SOI apporte donc dans cet exemple une fréquence fonctionnelle à 1,2V supérieure de 40% à celle du BULK et une consommation par opération inférieure de 25%.

#### 1.2.5.2. Portage direct d'un circuit de moyenne complexité

Nous considérons un bloc logique composé d'un nombre limité de cellules standards constituant un oscillateur en anneau commandé. Ce bloc logique est représentatif des performances et de la consommation dynamique d'une technologie. Le nombre de portes qui composent le circuit définit sa fréquence maximale. Les autres paramètres agissant sur la fréquence sont la température, la tension de seuil, les dimensions... mais surtout la technologie.

La figure montre la période minimale du circuit en fonction de sa consommation dynamique pour une tension de seuil basse, « LVt », conçu en BULK et porté directement en PD-SOI.

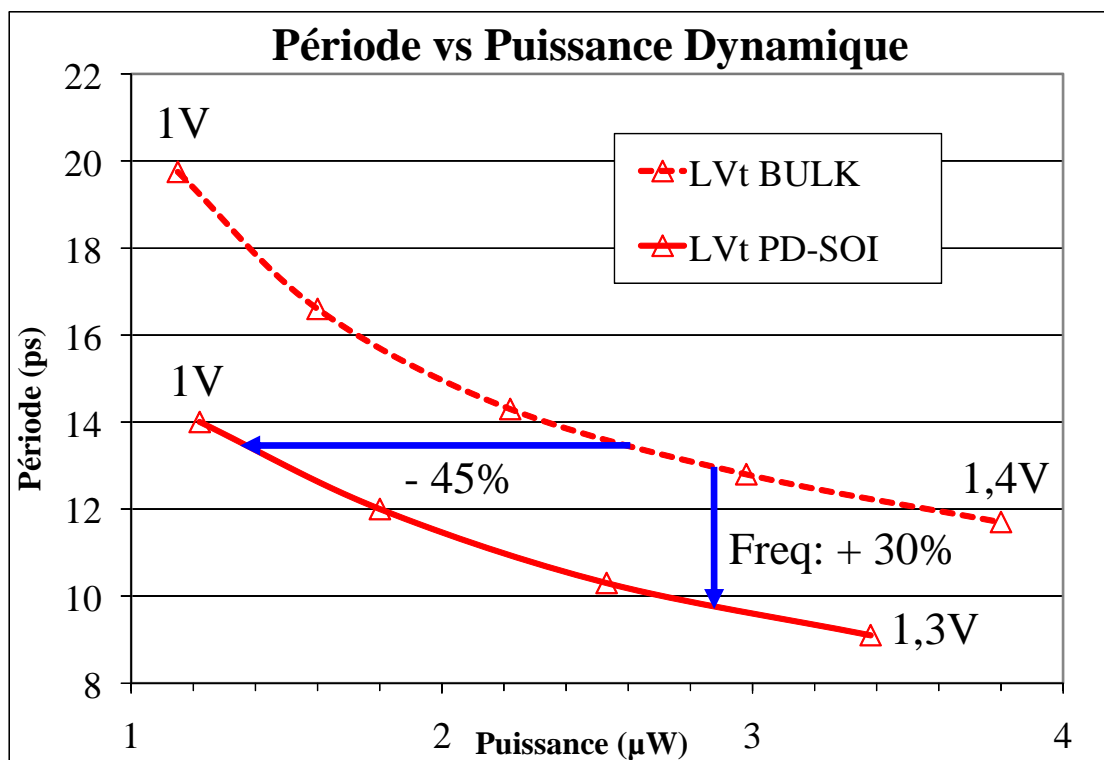


figure : Période en fonction de la Puissance Dynamique d'un circuit LVt en BULK et PD-SOI à 25°C en 65nm

Par rapport à la technologie BULK, deux comparaisons sont possibles. Dans le cas d'une utilisation haute performance, en travaillant avec une alimentation nominale ou au-delà,

on obtient une fréquence d'utilisation plus élevée de +35% à 1,2V en PD-SOI et BULK et de +30% pour la même consommation dynamique. L'autre cas correspond à des applications de basse consommation. Par rapport au BULK, pour une même fréquence de fonctionnement, un circuit porté de manière directe en technologie PD-SOI requiert une tension d'alimentation, inférieure de -18%, soit 1V en PD-SOI au lieu de 1,21V en BULK. Dans ce cas la consommation dynamique est plus faible de 45%.

### 1.2.5.3. Consommation Statique

A l'image du courant de fuite du MOSFET, la consommation statique d'un circuit est plus élevée en technologie PD-SOI qu'en technologie BULK, en raison de l'effet du body flottant. En reprenant le circuit précédant, alimenté mais à l'arrêt, la figure montre sa consommation statique en technologies BULK et PD-SOI. Comme précédemment, le circuit est de type LVt, mais une version de type HVt est présentée. La tension de seuil basse LVt permet d'avoir des performances élevées, alors que la tension de seuil haute HVt permet d'avoir une faible consommation statique.

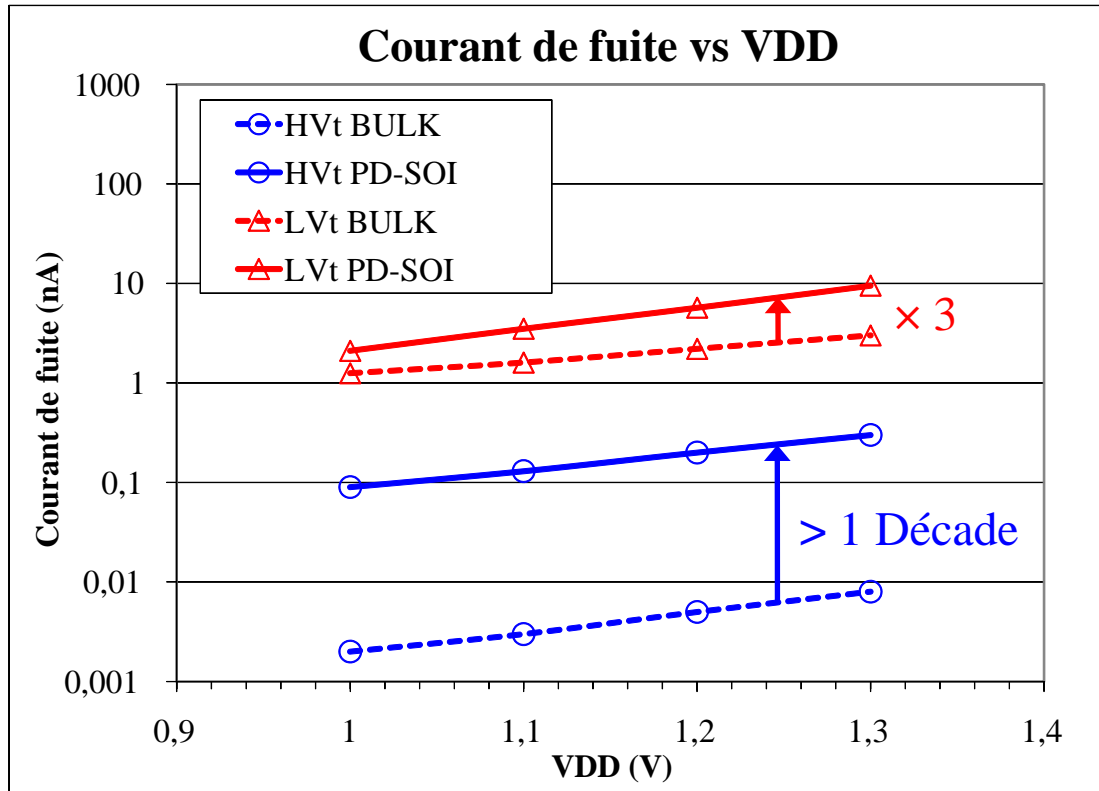


figure : Courant de fuite versus VDD d'un circuit LVt et HVt en BULK et PD-SOI à 25°C en 65nm.

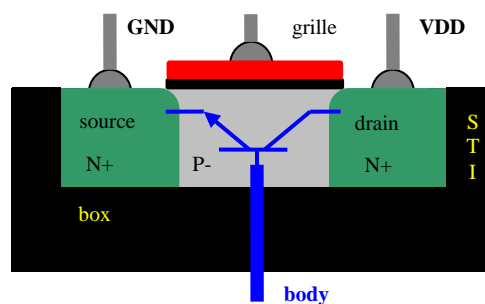
La figure montre que le courant de fuite est plus élevé en technologie PD-SOI. Il existe un facteur de 3 pour le circuit en LVt et plus d'une décade en HVt entre les fuites en PD-SOI et celles en BULK. L'effet du body flottant « FBE » est supérieur en HVt. Il désavantage la technologie PD-SOI en consommation statique par rapport au BULK.

Il existe plusieurs solutions à ce problème : la première est de procéder à la fabrication d'un transistor avec effet de body flottant réduit au maximum grâce à des implémentations spécifiques, notamment du drain et de la source [Hua'07]. Cependant, la solution la plus radicale est de fabriquer un transistor avec prise body qui permet de supprimer l'effet de body flottant grâce à un contrôle du potentiel de body  $V_b$  et donc de la tension de seuil  $V_{th}$ . Ce transistor est communément appelé « BC » pour Body Contacted en anglais.

### 1.3. Transistor avec prise body

#### 1.3.1. Inconvénient du transistor à body flottant (FB)

Le transistor FB présente de nombreux avantages, mais possède aussi certains inconvénients liés au body flottant comme l'augmentation de la consommation statique ou l'effet d'histoire. Or certaines applications nécessitent de faibles fuites comme les interrupteurs de puissance [Mut'95], et d'autres, comme les capacités variables [Che'03], ont besoin du contrôle du potentiel de body. Il en est de même pour les Transistors utilisés dans les protections contre les décharges électrostatiques [Vol'96], qui utilisent notamment l'effet bipolaire parallèle du MOS permettant de déclencher plus rapidement les protections (figure ).



**figure : PD-SOI MOS utilisé en Bipolaire parallèle**

C'est pour cela qu'un transistor à body contacté (BC) a été conçu pour contrôler le potentiel du body.

#### 1.3.2. Le transistor à body contacté



Le transistor BC est obtenu selon le même procédé de fabrication que le transistor FB. Il dépend d'une astuce « layout » d'implémentation qui n'a aucun sens en BULK. Ce type de transistor n'est disponible que dans la plateforme de conception PD-SOI.

L'oxyde latéral STI déposé autour de la zone active isole le body des transistors FB. Il suffit, pour créer un transistor BC, d'élargir la zone active du body et d'y inclure une prise dopée du même type [Mat'89]. La figure montre le principal moyen d'obtenir un transistor de type BC : en forme de T.

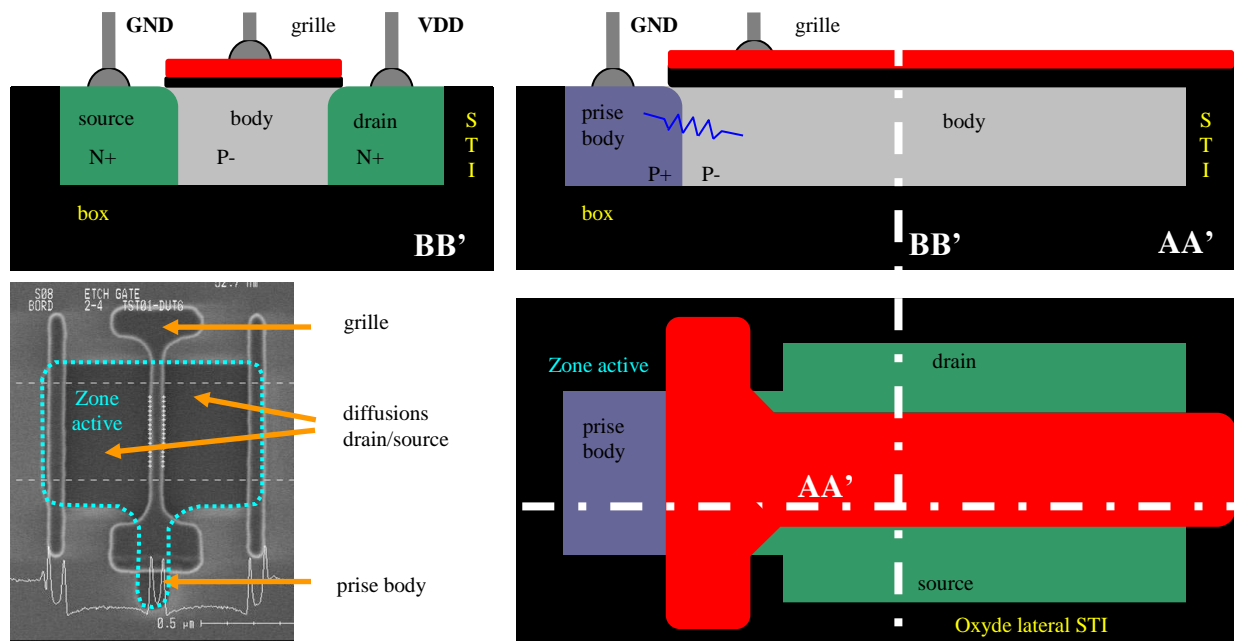


figure : BC MOSFET : vue layout

Il existe d'autres formes pour ces transistors [Cai'06], cependant les plus répandues restent les formes en « T » ou en « H », cette dernière n'étant que le « T » avec deux prises body, une sur chaque extrémité du MOSFET. L'accès au body sous le canal étant plus résistif en PD-SOI qu'en BULK en raison de la plus faible épaisseur de silicium, il est primordial pour des largeurs élevées de transistor de placer non plus une mais deux prises substrat de chaque côté du MOSFET et ainsi d'obtenir un MOSFET de forme « H » [All'91]. Il existe une largeur maximale par prise substrat, au-delà de laquelle un effet de body flottant peut apparaître à nouveau [Sch'03]. En raison de la présence de ces prises, on peut donc en déduire que la surface occupée est plus élevée pour des MOSFET de type BC que pour des MOSFET de type FB ou ceux en technologie BULK.

Le comportement électrique d'un MOSFET de type SOI BC est comparable à celui obtenu en technologie BULK [Pel'02] en termes de courants statiques  $I_{ON}$  et  $I_{OFF}$ . En revanche, l'ajout de polysilicium, permettant de créer la prise body, apporte une capacité de grille plus grande que le BULK. Le gain en performance obtenu grâce à la réduction des capacités de jonction  $C_{DS}$  peut être perdu avec le supplément de capacité de grille  $C_{GS}$  ajouté, comme l'illustre la figure . La source reste connectée au body pour le BULK et le PD-SOI.

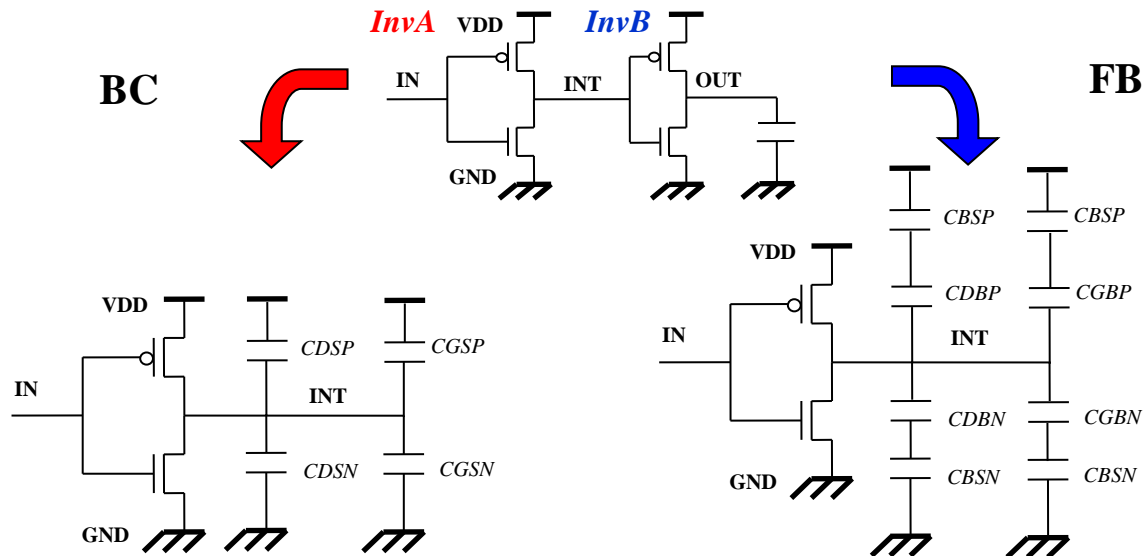


figure : Capacités vues par le nœud intermédiaire INT : modèle paramétrique au premier ordre.

De plus, les règles de conception des MOSFET BC n'autorisent pas une largeur minimale de grille  $W$  aussi petite que pour les MOSFET FB. En comparaison avec le BULK, une logique de type CMOS BC n'apporterait donc aucun gain en performances (tableau ), et introduirait même une perte en surface.

Deux inverseurs sont comparés, ils ont une largeur de PMOS de  $W=0,5\mu m$  et de NMOS de  $W=0,3\mu m$ , pour une tension d'alimentation de 1,2V à 25°C et une longueur de grille  $L=60nm$ . Le temps de propagation (50% de VDD) de l'inverseur implémenté avec des transistors FB est inférieur de 32% par rapport à celui implémenté avec des transistors BC.

Temps de propagation à 50% VDD : FB & BC		
	PD-SOI FB	PD-SOI BC
IN $\rightarrow$ OUT ( <i>InvA</i> $\rightarrow$ <i>InvB</i> )	21,7ps (-32%)	32,15ps

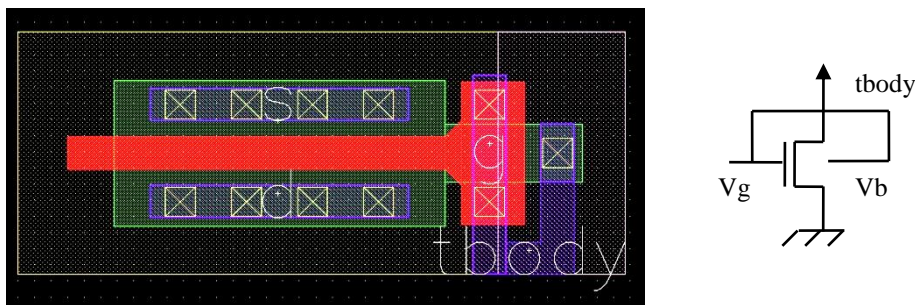
**tableau : Comparaison Body flottant (FB) versus Body contacté (BC) du temps de propagation à 50% de 2 inverseurs.**

C'est pour cela que les transistors de type BC ne sont utilisés dans un circuit que pour des applications de type « basse consommation », « suppression de l'effet d'hysteresis », « contrôle de la tension de seuil » et sont évités dans la logique.

### 1.3.3. Le transistor à body contacté connecté en DTMOS

Au lieu de relier le body du MOSFET BC à sa source, il est possible de le contrôler dynamiquement comme le propose le montage en DTMOS : body connecté à la grille.

La connexion d'un MOS BC en DTMOS (figure ), pour MOS à tension de seuil dynamiquement variable (Dynamic Threshold MOS [Ass'94]), permet d'augmenter la pente sous le seuil et donc le courant, ce qui a tendance à augmenter les performances.



**figure : NMOS BC configuré en DTMOS avec la grille connectée au body  $V_g = V_b$**

Le principe de celui-ci est de moduler favorablement la tension de seuil. Lorsque le transistor devient passant, il polarise le body de manière à réduire  $V_{th}$  et augmenter la vitesse. A l'opposé lorsque le transistor est bloqué, la polarisation du body augmente la tension de seuil  $V_{th}$  réduisant les fuites (cette modulation est mieux détaillée dans la partie concernant la réduction de la consommation traitant de la polarisation de body). Dans cette configuration, il peut fonctionner uniquement à une tension d'alimentation VDD inférieure à la tension de coude de la diode  $V_D$ , afin de garantir que la diode entre le body et la source ne soit jamais directe. Pour des tensions plus élevées, il faut introduire un limiteur un courant.

Ce type de transistor est possible uniquement en PD-SOI car les body des transistors restent isolés les uns des autres. En BULK, les body des transistors d'un même caisson (substrat) sont communs et par conséquent individuellement non polarisables.

### 1.4. Conclusion

Qu'ils soient issus d'un portage direct ou d'un flot de conception PD-SOI, les transistors de type FB permettent d'augmenter les performances et de diminuer la consommation dynamique des circuits par rapport à la technologie BULK. Les transistors de type BC permettent de réduire ou d'atteindre la même consommation statique que le BULK. Néanmoins ce transistor doit être implémenté de manière ponctuelle afin qu'il ne vienne ni détériorer les performances ni augmenter la surface.

- [Ada'01] Adan, A. & Higashi, K. "OFF-State leakage current mechanisms in bulkSi and SOI MOSFETs and their impact on CMOS ULSIs standby current" *IEEE Transactions on Electron Devices*, vol.48, p.2050-2057, 2001
- [Ade'07] Adell, P.; Barnaby, H.; Schrimpf, R. & Vermeire, B. "Band-to-Band Tunneling (BBT) Induced Leakage Current Enhancement in Irradiated Fully Depleted SOI Devices" *IEEE Transactions on Nuclear Science*, vol.54, p.2174-2180, 2007
- [Aga'05] Agarwal, A.; Mukhopadhyay, S.; Kim, C.; Raychowdhury, A. & Roy, K. "Leakage power analysis and reduction: models, estimation and tools" *IEEE Proceedings on Computers and Digital Techniques*, vol.152, p.353-368, 2005
- [All'91] Alles, M.; Kerns, S.; Massengill, L.; Clark, J.; Jones, K.L., J. & Lowther, R. "Body tie placement in CMOS/SOI digital circuits for transient radiation environments" *IEEE Transactions on Nuclear Science*, vol.38, p.1259-1264, 1991
- [Ass'94] Assaderaghi, F.; Parke, S.; Sinitsky, D.; Bokor, J.; Ko, P. & Hu, C. "A dynamic threshold voltage MOSFET (DTMOS) for very low voltage operation" *IEEE Electron Device Letters*, vol.15, p.510-512, 1994
- [Ban'97] Banna, S.; Chan, P. & Lau, J. "On buried oxide effects in SOI lateral bipolar transistors" *IEEE Transactions on Electron Devices*, vol.44, p.139-144, 1997
- [Ber'00] Bernstein, K. & Rohrer, N.J. "SOI Circuit Design Concepts" Kluwer Academic Publishers, Dordrecht, the Netherlands, 2000
- [Cai'06] Cai, X. & Hai, C. "Study of body contact of partial depleted SOI NMOS devices" *International Conference on Solid-State and Integrated Circuit Technology*, p.212-214, 2006
- [Cai'07] Cai, J.; Majumdar, A.; Dobuzinsky, D.; Ning, T.; Koester, S. & Haensch, W. "Ultra-Low Leakage Silicon-on-Insulator Technology for 65 nm Node and Beyond" *IEEE International Electron Devices Meeting*, p.907-910, 2007
- [Che'03] Chen, H.-Y.; Chen, K.-M.; Huang, G.-W.; Huang, C.-H.; Yang T.-H. & Chang, C.-Y. "Layout design of high-quality SOI varactor" *International symposium on VLSI Technology, Systems, and Applications*, p.273-275, 2003
- [Che'08] Chen, Q.; Goo, J.-S.; Ly, T.; Chandrasekaran, K.; Wu, Z.-Y.; Thuruthiyil, C. & Icel, A. "Off-state leakage current modeling in low-power/high-performance partially-depleted (PD) floating-body (FB) SOI MOSFETs" *International Conference on Solid-State and Integrated-Circuit Technology*, p.301-304, 2008
- [Col'85] Colinge, J.-P. "Transconductance of Silicon-on-insulator (SOI) MOSFET's" *IEEE Electron Device Letters*, vol.6, p.573-574, 1985
- [Col'94] Colinge, J.-P. "Recent advances in SOI technology" *International Electron Devices Meeting*, p.817-820, 1994
- [Col'04] Colinge, J.-P. "Silicon-On-Insulator technology: Materials to VLSI" Kluwer Academic Publishers, Dordrecht, the Netherlands, 2004

- [Del'09] Delprat, D.; Boedt, F.; David, C.; Reynaud, P.; Alami-Idrissi, A.; Landru, D.; Girard, C. & Maleville, C. "SOI substrate readiness for 22/20 nm and for fully depleted planar device architectures" *IEEE International SOI Conference*, p.1-4, 2009
- [Fla'10] Flatresse, P.; Le coz, J. "Partially Depleted SOI Design for Low-Power Applications" ESSDERC tutorial, 2010
- [Fos'98] Fossum, J.G.; Pelella, M. & Krishnan, S. "Scalable PD/SOI CMOS with Floating Bodies" *IEEE Electron Device Letters*, vol.19, p.414-416, 1998
- [Guo'06] Guo, D.; Bryant, A.; Wang, X.; Narasimha, S.; Miller, R. & Khare, M. "Gate-dielectric permittivity and Metal-gate work-function tradeoff in Lmet=25nm PDSOI device characteristics" *IEEE Electron Device Letters*, vol.27, p.505-507, 2006
- [Hua'07] Huang, R.; Chen, T.; Hong, S.; Lin, Y.; Tsai, T.; Liu, E.; Yang, C.; Hsieh, Y.; Huang, Y.; Pelloie, J.-L.; Tsai, C. & Ma, G. "Optimizing Floating Body Effect & AC performance in 65nm PD-SOI CMOS" *IEEE International SOI Conference*, p.107-108, 2007
- [Kea'07] Keating, M. et al. "Low Power Methodology Manual For System-on-Chip Design" Springer, 2007
- [Kim'09] Kim, K.; Kuang, J.; Gebara, F.; Ngo, H.; Chuang, C.-T. & Nowka, K. "TCAD/Physics-Based Analysis of High-Density Dual-BOX FD/SOI SRAM Cell With Improved Stability" *IEEE Transactions on Electron Devices*, vol.56, p.3033-3040, 2009
- [Kot'02] Kotani, N.; Ito, S.; Yasui, T.; Wada, A.; Yamaoka, T. & Hori, T. "Suppression of leakage current in SOI CMOS LSIs by using silicon-sidewall body-contact (SSBC) technology" *Symposium on VLSI Technology*, p.44-45, 2002
- [Kri'98] Krishnan, S. & Fossum, J. "Grasping SOI floating-body effects" *IEEE Circuits and Devices Magazine*, vol.14, p.32-37, 1998
- [Lho'06] L'Hostis, N.; Thomas, O.; Haendler, S.; Amara, A.; Flatresse, P. & Belleville, M. "Silicon characterization of standby leakage reduction techniques in a 0.13um Low Power Partially-Depleted Silicon-On-Insulator Technology" *IEEE International Conference on Integrated Circuit Design and Technology*, p.1-4, 2006
- [Lin'99] Lin, S. & Kuo, J. "Temperature-dependent kink effect model for partially-depleted SOI NMOS devices" *IEEE Transactions on Electron Devices*, vol.46, p.254-258, 1999
- [Lio'03] Liot, V. & Flatresse, P. "A new fast method to compute steady state in PD-SOI circuits and its application to standard cells library characterization" *IEEE International SOI Conference*, p.170-171, 2003
- [Mat'89] Matloubian, M. "Smart body contact for SOI MOSFETs" *IEEE SOS/SOI Technology Conference*, p.128-129, 1989

- [Mat'01] Mathew, S. "Sub-500ps 64b ALUs in 0.18um SOI/Bulk CMOS: Design & Scaling Trends" *IEEE International Solid State Circuit Conference*, p.318-319, 2001
- [Mis'08] Mishra, U. & Singh, J. "Semiconductor Device Physics and Design" Springer, 2008
- [Mut'95] Mutoh, S.; Douseki, T.; Matsuya, Y.; Aoki, T.; Shigematsu, S. & Yamada, J. "1-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS" *IEEE Journal of Solid-State Circuits*, vol.30, p.847-854, 1995
- [Nar'01] Narendra, S.; Tschanz, J.; Keshavarzi, A.; Borkar, S. & De, V. "Comparative performance, leakage power and switching power of circuits in 150 nm PD-SOI and bulk technologies including impact of SOI history effect" *Symposium on VLSI Circuits*, p.217-218, 2001
- [Pel'02] Pelella, M. & Fossum, J. "On the performance advantage of PD/SOI CMOS with floating bodies" *IEEE Transactions on Electron Devices*, vol.49, p.96-104, 2002
- [Pel'03] Pelloie, J.-L. "Reliability issues in SOI technologies and circuits" *Bipolar/BiCMOS Circuits and Technology Meeting*, p.151-155, 2003
- [Pel'05] Pelloie, J.-L. "Using SOI to achieve low-power consumption in digital" *IEEE International SOI Conference*, p.14-17, 2005
- [Raj'95] Rajgopal, R.; Schiebel, R.; Iyer, S.; Joyner, K. & Houston, T. "Silicon film thickness and material dependence of "reverse short channel effect" for SOI NMOSFETs" *International Electron Devices Meeting*, p.533-536, 1995
- [Roy'03] Roy, K.; Mukhopadhyay, S. & Mahmoodi-Meimand, H. "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits" *Proceedings of the IEEE*, vol.91, p.305-327, 2003
- [Sak'06] Sakurai, T.; Matsuzawa, A. & Douseki, T. "Fully-Depleted SOI CMOS Circuit and Technology For Ultra Low-Power Applications" Springer, 2006
- [Sch'03] Schwank, J.; Ferlet-Cavrois, V.; Shaneyfelt, M.; Paillet, P. & Dodd, P. "Radiation effects in SOI technologies" *IEEE Transactions on Nuclear Science*, vol.50, p.522-538, 2003
- [Sin'98] Sinitsky, D.; Fung, S.; Tang, S.; Su, P.; Chan, M.; Ko, P. & Hu, C. "A dynamic depletion SOI MOSFET model for SPICE" *Symposium on VLSI Technology*, p.114-115, 1998
- [Su'01] Su, P.; Goto, K.; Sugii, T. & Hu, C. "Self-heating enhanced impact ionization in SOI MOSFETs" *IEEE International SOI Conference*, p.31-32, 2001
- [Suz'00] Suzuki, K. "Short channel MOSFET model using a universal channel depletion width parameter" *IEEE Transactions on Electron Devices*, vol.47, p.1202-1208, 2000

- [Vol'96] Voldman, S.; Schulz, R.; Howard, J.; Gross, V.; Wu, S.; Yapsir, A.; Sadana, D.; Hovel, H.; Walker, J.; Assaderaghi, F.; Chen, B.; Sun, J.Y.-C. & Shahidi, G. "CMOS-on-SOI ESD protection networks" Electrical Overstress/Electrostatic Discharge Symposium, p.291-301, 1996
- [Wei'98] Wei, A.; Sherony, M. & Antoniadis, D. "Effect of floating-body charge on SOI MOSFET design" *IEEE Transactions on Electron Devices*, vol.45, p.430-438, 1998
- [Yeh'02] Yeh, W.-K.; Wang, W.-H.; Fang, Y.-K.; Chen, M.-C. & Yang, F.-L. "Hot-carrier-induced degradation for partially depleted SOI 0.25-0.1  $\mu\text{m}$  CMOSFET with 2-nm thin gate oxide" *IEEE Transactions on Electron Devices*, vol.49, p.2157-2162, 2002





## 2. Les techniques de réduction de la consommation

*Ce chapitre aborde la question de la consommation énergétique des circuits en technologies sur substrat massif BULK et en technologie PD-SOI. Les principales techniques de réduction de la consommation, aussi bien dynamique que statique, sont présentées indépendamment des technologies. La technique de réduction de la consommation statique MTCMOS, la plus efficace et la mieux adaptée à la technologie PD-SOI, basée sur les interrupteurs de puissance est introduite dans ce chapitre*

## 2.1. Puissance dissipée dans les Circuits intégrés numériques

### 2.1.1. Puissance dynamique et statique

La puissance totale se définit comme la puissance dissipée nécessaire au traitement et au maintien de l'information. La puissance liée au traitement de l'information est la puissance dynamique alors que la puissance nécessaire au maintien de l'information est la puissance statique [Kea'07].

La puissance totale  $P_{total}$  d'un circuit numérique CMOS peut être modélisée par l'équation [Cha'92]. Le premier terme correspond à la puissance dynamique et le second terme correspond à la puissance statique.  $p_t$  est la probabilité qu'une porte commute - toutes les portes logiques ne commutent pas à chaque cycle d'horloge -  $C_{OUT}$  est la capacité de chargement,  $VDD$  la tension d'alimentation,  $F_{CLK}$  la fréquence,  $t_{sc}$  et  $I_{sc}$  respectivement le temps et le courant de court-circuit et  $I_{DDQ}$  le courant de fuite.

$$P_{total} = P_{dynamique} + P_{statique}$$

$$P_{dynamique} = p_t (C_{OUT} \times VDD^2 \times F_{CLK} + t_{sc} \times I_{sc} \times VDD)$$

$$P_{statique} = I_{DDQ} \times VDD$$

équation

Il faut bien distinguer la différence entre énergie et puissance. L'énergie se caractérise par le produit de la puissance moyenne dissipée avec le temps.

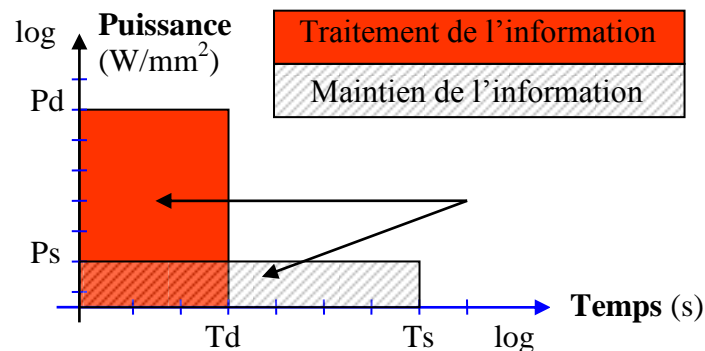


figure : Energie versus Puissance

Comme le montre la figure, dans le cas du traitement de l'information, l'énergie est égale au temps durant lequel le circuit est actif «  $T_d$  » multiplié par la puissance moyenne

dissipée « Pd ». Similairement, l'énergie de maintien est le produit de la puissance statique « Ps » par le temps durant lequel le circuit est inactif « Ts ». Le temps actif Td est variable : c'est le temps pour effectuer une ou plusieurs successions de tâches, il dépend de la fréquence et du nombre de cycles. Ces paramètres vont dépendre du type d'application visée.

### 2.1.2. Comparaison des besoins des applications fixes et portables

Ce qui distingue deux applications, c'est le ratio entre le temps de veille et le temps d'activité. Ainsi Ts est proche de Td pour les applications fixes et Ts est bien supérieur à Td pour des applications mobiles. Le besoin d'autonomie, dans ce second cas, impose une gestion de l'énergie et l'utilisation de circuits de technologie faible consommation.

En effet, les applications portables sont alimentées sur batterie, et pour des raisons d'autonomie, ces applications doivent être peu gourmandes en énergie. Ce qui n'est pas toujours le cas pour les circuits d'application fixe où les performances sont plutôt recherchées et où l'alimentation est branchée sur le réseau électrique.

$$Energie_{totale} = P_{dynamique} \times Td + P_{statique} \times Ts$$

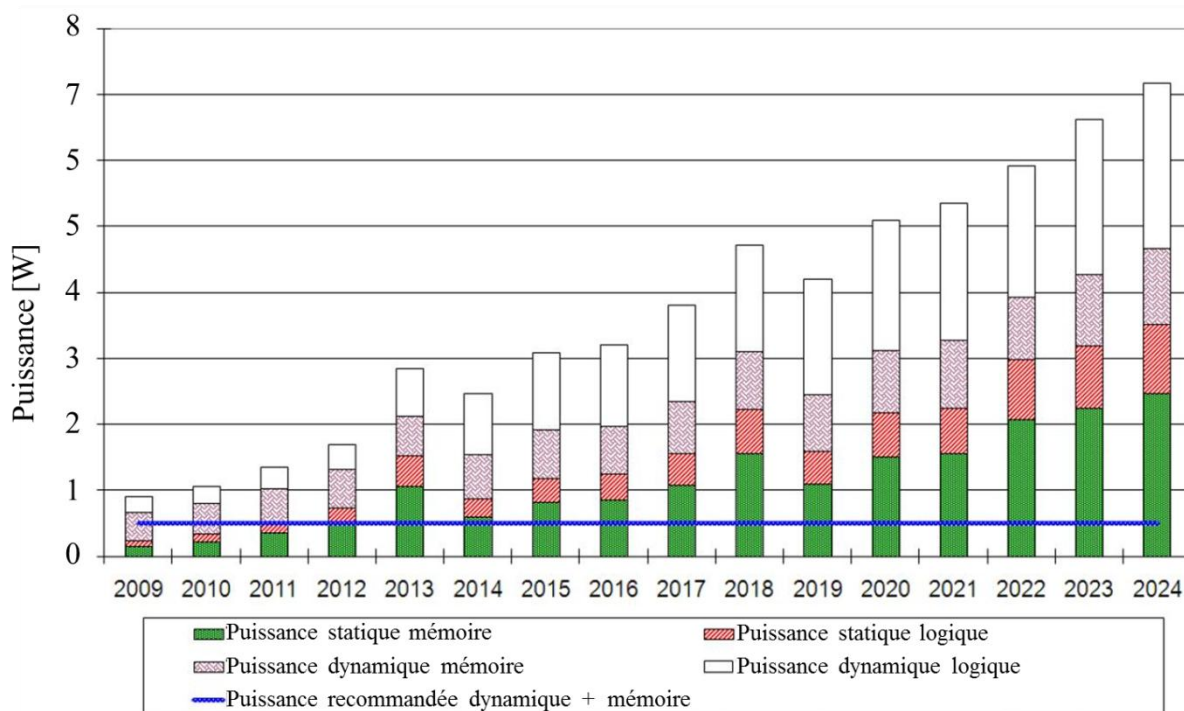
**équation**

L'équation représente l'énergie totale en fonction de la puissance dynamique et statique, et en fonction des temps d'activité et de veille. Pour une application de type fixe, il faut se pencher principalement sur la réduction de la consommation dynamique. Dans le cas des applications mobiles ou portables, il faut également inclure une réduction de la consommation statique.

### 2.1.3. Evolution avec les technologies

Jusqu'au nœud technologique 90nm, la consommation électrique d'un circuit en fonctionnement était majoritairement liée à la puissance dynamique. Il suffisait dès lors de réduire la puissance dynamique pour réduire la puissance totale. Ce n'est plus le cas pour les technologies inférieures à 90nm, développées depuis 2007. La puissance statique devient une

composante de plus en plus importante, comme le montre la projection de l'ITRS « International Technology Roadmap for Semiconductor » mise à jour en 2010, figure .



**figure : Puissance totale prévue par l'ITRS en 2010 pour des applications mobiles**

Il est clair que le ratio entre la puissance dynamique et statique devient de plus en plus faible. Il est donc logique d'aborder la réduction de la consommation électrique totale d'un circuit autrement, en prenant de plus en plus en considération la composante statique. Que ce soit en technologie BULK ou en PD-SOI, la réduction de la consommation électrique totale est un challenge.

L'évolution des technologies a engendré une réduction de la tension d'alimentation VDD, à la fois pour ne pas dégrader la fiabilité des transistors, mais aussi pour réduire la puissance dynamique. Afin d'augmenter les performances, la tension de seuil,  $V_{th}$ , a également été réduite dans le but de maintenir un rapport  $VDD/V_{th}$  constant [Van'04].

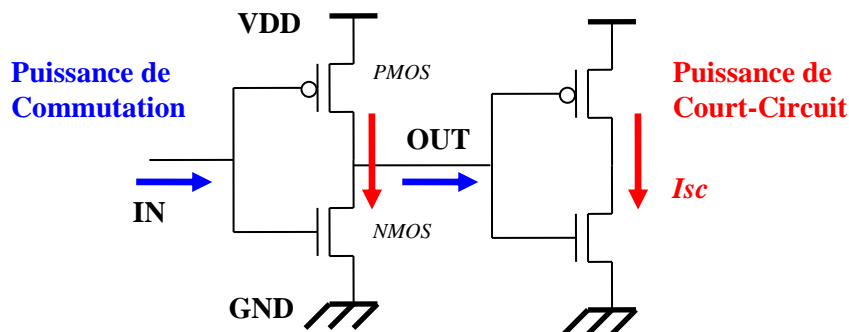
Dans la suite de ce chapitre nous allons évoquer les principales techniques qui permettent de réduire la consommation électrique, la consommation dynamique dans un premier temps, puis la consommation statique dans un second temps. L'objectif est de retenir la technique la plus avantageuse pour la technologie 65nm PD-SOI sachant que son principal inconvénient est qu'elle fuit plus que la technologie BULK.

## 2.2. Techniques de réduction de la puissance dynamique

### 2.2.1. La puissance dynamique

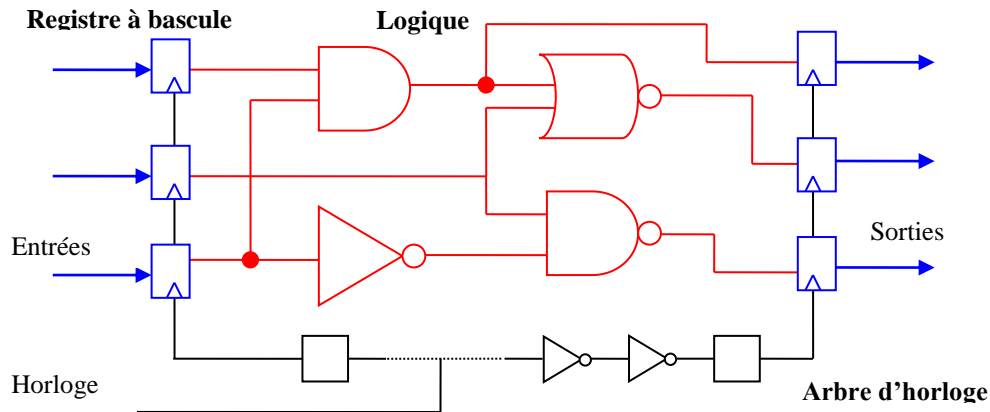
La puissance dynamique, présentée dans cette partie, est la composante dominante de la consommation électrique totale d'un circuit. La réduire, c'est assurer un abaissement indispensable de la consommation totale.

En reprenant l'équation, le premier terme de la puissance dynamique est la puissance liée au chargement des capacités de grille des transistors des portes logiques et des capacités équivalentes des interconnexions, comme le montre la figure. Le second terme correspond à la puissance de court-circuit : l'entrée d'une porte logique est connectée aux grilles de transistors PMOS et NMOS. Lorsque le potentiel de cette entrée commute, son passage de 0 à VDD entraîne aux alentours de  $VDD/2$  l'activation simultanée des deux transistors PMOS et NMOS créant ainsi, durant un court instant, un court-circuit,  $I_{sc}$ , entre l'alimentation VDD et la masse GND (figure). Ce courant reste négligeable par rapport au courant de commutation, pour des temps de transition en entrée classiquement trouvés dans les circuits.



**figure : Puissance dynamique d'une porte logique**

Cette consommation électrique dépend donc de l'activité, de l'alimentation, du nombre de transistors et de la fréquence de fonctionnement. Dans un circuit logique numérique, le traitement des données est rythmé, via les bascules, par la fréquence d'horloge  $F_{CLK}$  (figure). Etant donné le grand nombre de bascules, la distribution de l'horloge, nommée « arbre d'horloge », est propagée grâce à un chemin constitué d'inverseurs et de suiveurs. La puissance dynamique est donc la somme de la puissance dynamique du cœur logique, des bascules, de l'arbre d'horloge mais également des mémoires, lors de l'écriture et de la lecture.



**figure : Vue schématique d'un circuit numérique synchronisé par l'horloge**

Une fois la consommation dynamique définie et posée sous forme d'équation (équation ), il reste à présenter les techniques qui permettent sa réduction.

$$P_{dynamique} = p_t (C_{OUT} \times VDD^2 \times F_{CLK} + t_{sc} \times I_{sc} \times VDD)$$

**équation**

### 2.2.2. Réduction de la puissance dynamique

Dans cette partie, les techniques réduisant la consommation dynamique sont introduites et détaillées. Ces techniques vont de la réduction de la tension d'alimentation au changement de fréquence en passant par l'optimisation en termes d'architecture et par l'interruption des données et de l'horloge

#### 2.2.2.1. Réduction de l'alimentation VDD

Au vue de la formule de la puissance dynamique (équation ), le premier paramètre qui permet la réduction de la puissance dynamique est la tension d'alimentation VDD. La réduction de la tension d'alimentation VDD s'appelle le « voltage scaling ». Cela permet une réduction de la consommation proportionnelle à  $V^2$ . Il est important de souligner que la fréquence d'horloge fonctionnelle maximale  $F_{max}$  d'un circuit dépend de VDD. En effet,  $F_{max}$  dépend du courant  $I_{ds}$  qui fait commuter les portes logiques et qui est fonction de VDD et de  $V_{th}$  (équation ) :  $W$  est la largeur du NMOS,  $L$  la longueur de grille,  $\mu_n$  la mobilité des

électrons,  $C_{ox}$  la capacité de grille,  $V_{gs}$  la tension grille-source,  $V_{ds}$  la tension drain-source et  $V_{th}$  la tension de seuil. Lorsque le NMOS est passant  $V_{gs}=V_{DD}$ .

$$I_{ds} = \frac{W}{L} * \mu_n C_{ox} * \left( V_{gs} - V_{th} - \frac{V_{ds}}{2} \right) * V_{ds}$$

équation

Il est possible de réduire  $V_{DD}$  lorsque  $F_{clk}$  est en dessous de la fréquence maximale fonctionnelle  $F_{max}$ . Le tableau résume, à titre d'exemple, les différentes fréquences maximales fonctionnelles d'un bloc logique en fonction de son alimentation  $V_{DD}$ .

VDD (V)	Fréquence maximale $F_{max}$ (MHz)
1	250
1,1	310
1,2	340

**tableau : Fréquence maximale d'un bloc logique en fonction de l'alimentation  $V_{DD}$**

On peut constater que si la fréquence de fonctionnement visée  $F_{clk}$  est de 300Mhz, il est possible de réduire la tension de 1,2V à 1,1V sans aucun impact sur les performances. La fréquence maximale dépend des paramètres extérieurs comme la température mais également des paramètres de fabrication, tels que le dopage de canal ou la longueur de grille des MOSFET...

La technique appelée « dynamic voltage scaling » permet de réduire dynamiquement le potentiel de l'alimentation  $V_{DD}$  et donc la consommation dynamique. Dans certains cas, la variabilité des procédés et le changement de température entraînent une fluctuation des paramètres tels qu'une réduction de la tension de seuil  $V_{th}$ . Dans le but de maintenir un ratio  $V_{DD}/V_{th}$  constant, la technique « dynamic voltage scaling » permet de compenser un abaissement de la tension de seuil en réduisant dynamiquement le potentiel d'alimentation  $V_{DD}$  [Mou'10].

#### 2.2.2.2. Changement de fréquence

La réduction de la fréquence « frequency scaling » a pour effet de ralentir un circuit qui a des besoins moindres en performance à des instants différents. Par exemple, un



ordinateur a besoin de hautes performances pour les jeux vidéo et de performances moins élevées pour le traitement de texte. La réduction de la fréquence ne change en rien l'énergie nécessaire au traitement de l'information si elle n'est pas accompagnée d'un « voltage scaling » [Lin'09]. En effet, à même tension d'alimentation, diminuer la fréquence permet uniquement de réduire la puissance dissipée moyenne mais rallonge le temps de calcul, ce qui ramène au final à la même énergie dissipée. Le bilan énergétique d'une opération logique ne dépend donc pas de la fréquence mais du nombre de cycle dont elle a besoin pour sa complète exécution, et de sa tension d'alimentation.

### 2.2.2.3. Optimisation de l'architecture

Certaines méthodes visent à réduire au maximum le nombre de transistors qui vont commuter physiquement ou temporellement, ce qui va réduire la capacité globale équivalente du circuit  $C_{OUT}$  ou la probabilité de commutation  $p_t$  (équation ). Afin d'obtenir ce résultat, l'architecture d'un circuit peut être choisie de manière à plus ou moins sérialiser ou paralléliser le calcul. Ce choix est déterminé par la puissance maximale disponible, les fréquences à atteindre et la surface de silicium disponible. Cette étape est réalisée par les outils de conception. Une optimisation de la consommation dynamique peut donc introduire une dégradation des performances, de la fonctionnalité ou de la surface.

### 2.2.2.4. Contrôle des données et de l'horloge

L'interruption du flux de données en entrées permet de réduire au maximum la probabilité  $p_t$  de l'équation . Le fait que les entrées d'un circuit ne changent plus, met le circuit dans un mode de maintien ou de non-commutation. Ce mode permet à la partie logique de ne plus consommer dynamiquement. Il ne reste plus que la puissance statique et la puissance de l'arbre d'horloge et des registres d'entrées/sorties.

L'interruption de l'horloge, « clock gating » permet d'obtenir un arrêt complet de la partie logique et, de surcroît, l'extinction complète de l'arbre d'horloge [Mah'09]. Le « clock gating » permet de tout interrompre, ceci d'une manière simple et sans aucune complexité, réduisant ainsi au maximum la puissance dynamique. L'association d'un bloc gérant le « clock gating » et l'interruption des entrées permet de totalement couper la consommation dynamique du circuit.

## 2.2.2.5. Point de fonctionnement optimal

La puissance dynamique dépend de VDD de manière quadratique, puisqu'elle est proportionnelle à  $CV^2f$ . La réduction simultanée de VDD et de  $V_{th}$  apporte une forte réduction de la puissance dynamique, tout en maintenant la possibilité de maintenir ou d'augmenter la fréquence maximale d'un circuit [Hor'94]. En effet, le délai d'inversion d'une porte logique  $T_p$  est fonction de VDD et de  $V_{th}$  comme le montre l'équation .

$$T_p = \alpha \frac{C_{OUT} \times VDD}{(VDD - V_{th})^2}$$

équation

La réduction simultanée de VDD et de  $V_{th}$  est la technique qui a le plus d'impact sur la réduction de la puissance totale. Malheureusement, la réduction de  $V_{th}$  implique une augmentation exponentielle des courants de fuite et donc de la puissance statique. La figure montre la puissance totale d'un oscillateur en anneau en technologie 0,13 $\mu$ m dont la fréquence reste inchangée sur une gamme de tension d'alimentation VDD allant de 400mV à 1,4V. Pour maintenir une fréquence d'oscillation constante, il suffit de moduler la tension de seuil  $V_{th}$  en fonction de VDD.

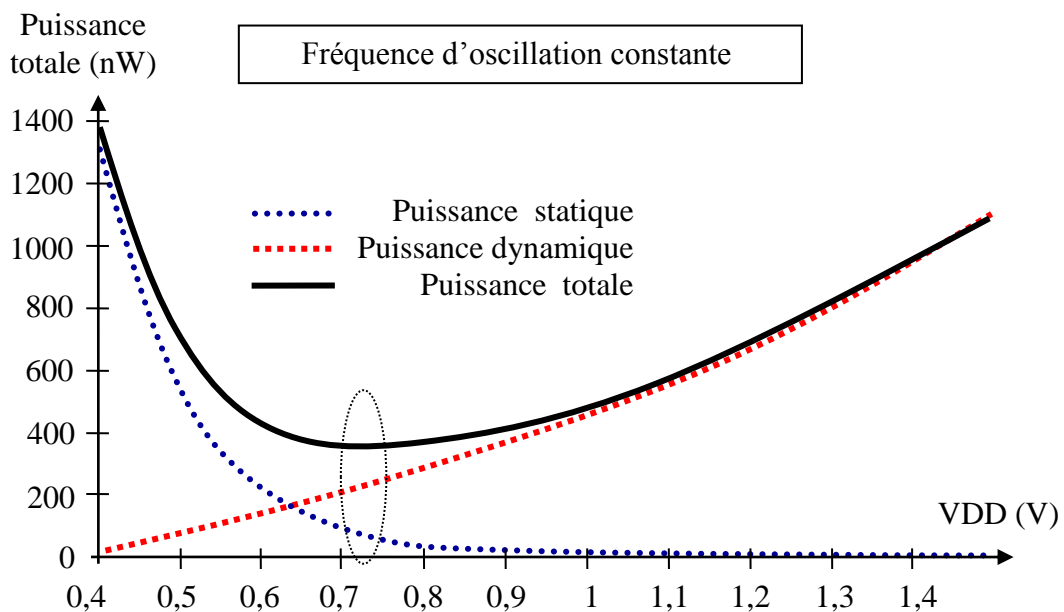


figure : Consommation électrique d'un oscillateur fonctionnant à la même fréquence

Il existe un optimum de tension d'alimentation VDD et Vth. En-dessous de cet optimum, la réduction de la puissance dynamique devient plus faible que l'augmentation de la puissance statique [Cal'04]. Cet optimum dépend aussi de l'activité du circuit [Pig'04].

Pour les technologies submicroniques, la puissance statique devient de plus en plus importante. Il faut donc basculer dans un mode de conception de circuit introduisant des techniques de réduction de la consommation statique. Dans la partie suivante, cette consommation ainsi que les diverses techniques pour la réduire y sont présentées.

### 2.3. Techniques de réduction de la puissance statique

#### 2.3.1. La puissance statique

L'augmentation de la puissance statique est liée principalement à la réduction de la tension de seuil Vth. Cette réduction est obtenue par une augmentation du dopage du canal, mais aussi grâce à la réduction de la distance entre drain et source. La puissance statique dépend principalement du courant de fuite sous le seuil du canal  $I_{DS}$  en fonction de la tension de seuil Vth (équation ) selon [Aga'05]. D'autres courants de fuite contribuent à l'augmentation de la consommation, néanmoins leurs impacts restent moindre comparés au courant de fuite du canal sous le seuil  $I_{DS}$  pour  $V_{gs}=0$ . De nouveaux paramètres rentrent en jeu dans l'équation du courant sous le seuil en plus de ceux présentés équation :  $q$  est la charge d'un électron,  $kT/q$  la tension thermique.

$$I_{DS} = \mu_n \times Cox \times \left( \frac{kT}{q} \right)^2 \times \frac{W}{L} \times e^{\frac{V_{gs}-V_{th}}{n \times kT/q}} \times \left( 1 - e^{\frac{-V_{ds}}{kT/q}} \right)$$

équation

Le courant de conduction sous le seuil  $I_{DS}$  croît de manière exponentielle avec la réduction de la longueur de grille des transistors. Cela provoque une augmentation de la puissance statique avec l'évolution des nœuds technologiques, malgré la diminution de la tension d'alimentation VDD. Comme le souligne la figure , la puissance statique devenant aussi importante que la puissance dynamique dans les technologies submicroniques, il devient

donc important pour les concepteurs de circuits intégrés de faible consommation de développer aussi bien des techniques de réduction de la consommation dynamique que statique.

### 2.3.2. Réduction de la puissance statique

Il existe plusieurs techniques de réduction de la puissance statique [Kea'07]. Parmi celles-ci, il est possible de citer la réduction de VDD, l'empilement de transistors, l'utilisation de transistors à longueur de grille non minimale, la polarisation de substrat, le multi-Vt. Ces techniques restent aujourd'hui les principales et plus utilisées.

Dans un niveau d'intégration plus haut, l'architecture d'un circuit numérique compte également, ainsi que le vecteur optimal appliqué aux entrées car le courant de fuite dépend de l'état des transistors.

#### 2.3.2.1. Réduction de l'alimentation VDD

Au même titre que la réduction de l'alimentation mène à la réduction de la puissance dynamique, elle agit également sur la réduction de la puissance statique comme le souligne l'équation .

Les techniques « voltage scaling » et « dynamic voltage scaling » s'appliquent donc pour la réduction de la consommation statique. Il est même envisageable de réduire encore plus la tension d'alimentation dans un mode de rétention lorsque le circuit est en veille ou soumis à une technique du type « clock gating ». Dans ce cas, la tension d'alimentation peut être abaissée jusqu'à la tension minimale de maintien de l'information dans les registres et mémoires.

## 2.3.2.2. Empilement de transistors

La technique d'empilement [Nar'01] consiste à placer non plus un transistor entre l'alimentation VDD ou GND et la sortie mais plusieurs comme l'illustre la figure .

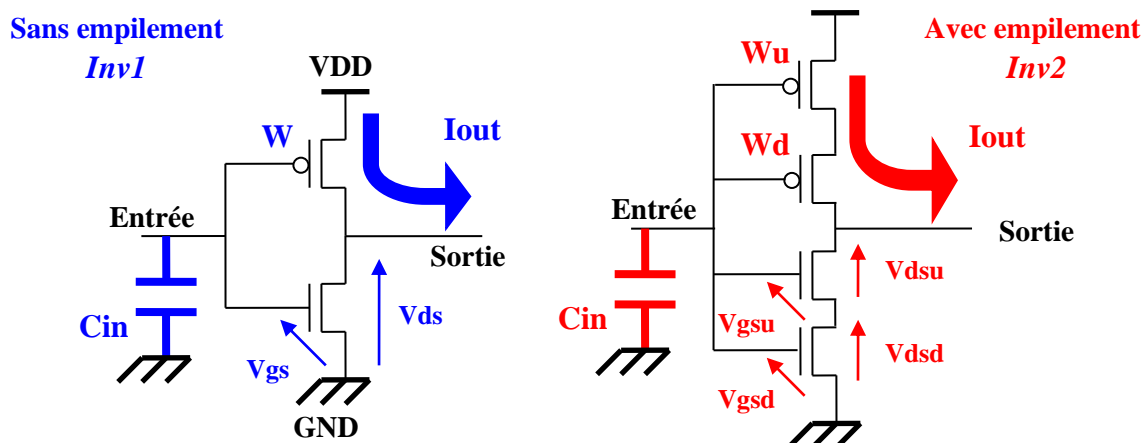
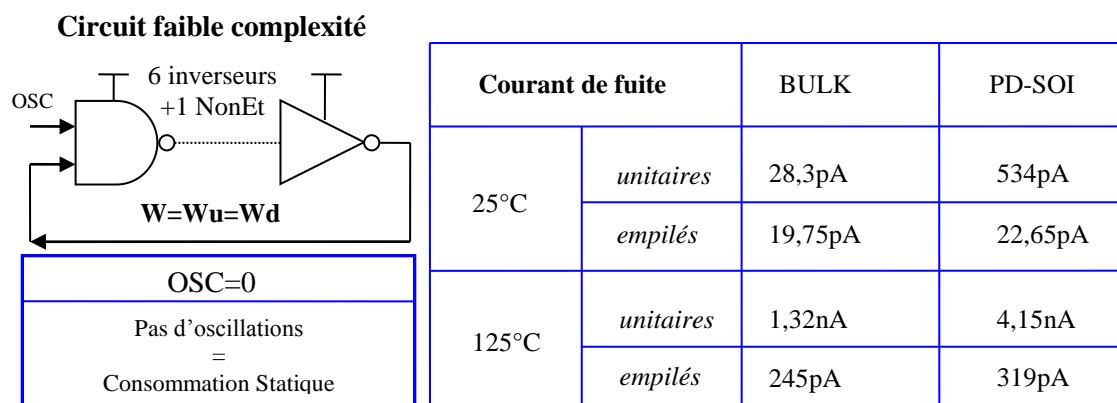


figure : Technique d'empilement appliquée à un inverseur.

Lorsque l'entrée est à la masse, la sortie de l'inverseur (*Inv1*) est à VDD induisant une différence de potentiel entre le drain et la source du NMOS,  $V_{ds}=V_{DD}$ . Suivant la même configuration, l'empilement de transistors NMOS aura pour conséquence de réduire la chute de potentiel  $V_{ds}$  de chacun des transistors empilés (*Inv2*) aux alentours de  $V_{DD}/2$  ce qui a pour effet de polariser négativement la grille du transistor du haut de la pile,  $V_{gsu}<0$ . Cette réduction de  $V_{ds}$  et de  $V_{gs}$  agit fortement sur la réduction du courant de canal sous le seuil. En 65nm PD-SOI, l'empilement de transistors est encore plus bénéfique qu'en CMOS BULK, comme le montre la figure . En effet la réduction de  $V_{ds}$  conduit à la réduction, voire même à la disparition, de l'effet de body flottant pour  $V_{ds} < 0,8V$ , annihilant ainsi l'accumulation de charges dans le body et la réduction de  $V_{th}$  responsable de l'accroissement de la consommation statique. De plus, le potentiel de body suit le potentiel de la source de chaque MOS, ce qui n'est pas le cas en BULK. Le fait que le potentiel de body suive le potentiel de source donne au transistor de meilleures performances en comparaison avec le BULK ; c'est d'ailleurs l'un des principaux intérêts de cet empilement.



**figure : Courant de fuite d'un circuit de faible complexité contenant des portes logiques avec MOS unitaires et empilés en 65nm BULK & PD-SOI LP HVt à 25°C**

A 25°C, la puissance statique du PD-SOI est 20 fois supérieure à celle du BULK avec des portes logiques composées de MOS unitaires contre seulement 15% supérieure avec des MOS empilés. L'inconvénient d'une telle technique est la perte en performances. En effet pour le même courant d'un inverseur,  $I_{out}$ , il faut multiplier par 2 la largeur de chacun des transistors NMOS et PMOS,  $W_u = W_d = 2 \times W$  ce qui a pour effet de multiplier par 4 la capacité d'entrée,  $C_{in}$ . A l'opposé pour la même capacité d'entrée de l'inverseur,  $C_{in}$ , il faudrait diviser par 2 la largeur des transistors NMOS et PMOS,  $W_u = W_d = W/2$ , ce qui a pour conséquence de diviser par 4 le courant de sortie utile à la commutation des étages logiques suivants,  $I_{out}$ . Il est donc conseillé d'utiliser cette technique là où la réduction des courants de fuite prime sur les performances. En d'autres termes, cette technique est bénéfique pour des circuits sans contraintes surfaciques où la fréquence est peu élevée et le temps de veille  $T_s$  est largement supérieur à  $T_d$ , le temps actif.

#### 2.3.2.3. Vecteur optimal

La technique d'empilement est naturellement observable dans un circuit pour des portes logiques à plusieurs entrées, telles que la porte NonEt lorsque ses entrées sont à GND ou la porte NonOu lorsque ses entrées sont à VDD. Pour une certaine configuration, un circuit logique aura donc un nombre plus ou moins important de transistors bloqués et empilés. Il existe donc un vecteur d'entrée appelé « vecteur optimal » qui amène une consommation statique minimale car c'est avec ce vecteur d'entrée que le circuit aura le plus de MOS bloqués et empilés en même temps. Il existe plusieurs méthodes pour obtenir ce vecteur, et il est même possible d'optimiser les transistors non empilés [Joh'02]. Aujourd'hui cette solution reste contraignante en termes d'outils de conception et de temps.

#### 2.3.2.4. Longueur de grille

L'évolution des technologies et la réduction de la longueur de grille entraîne une réduction du canal [Suz'00] et donc une augmentation du courant de fuite de canal. Une solution, le multi Lg [Cla'04], est d'introduire des transistors ayant des longueurs de grille différentes. Les transistors à longueur de grille non-minimale vont être utilisés dans un bloc logique pour réduire la consommation statique. En effet ces transistors sont connus pour leur plus faible puissance statique ; ils n'ont pas, ou peu, d'effet de canal court par rapport aux transistors avec une longueur de grille minimale. Mais comme ils ralentissent le bloc logique dans lequel ils sont insérés, il convient de remplacer ceux qui sont inclus dans les différents chemins critiques en délai, par des transistors à longueur de grille minimale, qui eux fuient plus mais vont beaucoup plus vite. Cette technique permet au bloc logique d'atteindre des fréquences plus élevées.

#### 2.3.2.5. Polarisation de Substrat

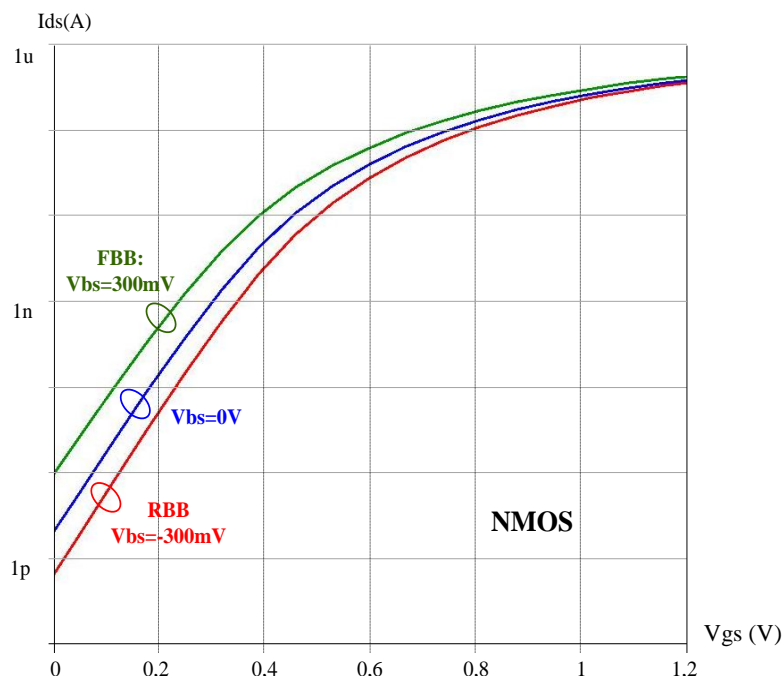
La polarisation du substrat ou de caisson dans le cas du BULK et de body dans le cas d'un transistor individuel en PD-SOI, est une technique plus flexible car elle peut être définie par l'utilisateur de manière temporelle, alors que les techniques d'empilement ou de multi Lg sont des techniques de conception par définition figées. La polarisation de body a pour objectif de faire varier la valeur de  $V_{th}$  en fonction du facteur de body  $\beta$ , équation . Ce qui a

pour conséquence de moduler le courant  $I_{ds}$  [Im'03].  $V_{th0}$  est la tension de seuil lorsqu'aucune polarisation de substrat ou body n'est appliquée.

$$V_{th} = V_{th0} - \beta \times \sqrt{V_{bs}}$$

équation

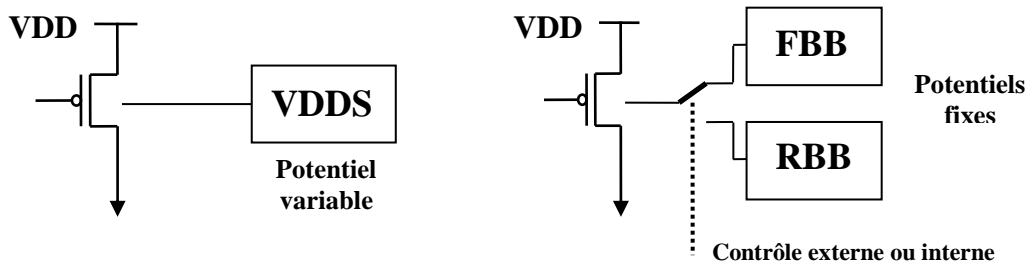
Cette technique aussi appelée VTCMOS pour « Variable Threshold CMOS » est largement utilisée par les concepteurs de circuits intégrés. Elle se décline sous plusieurs formes. La première forme vise à amener, au plus bas la valeur de  $V_{th}$  lorsque le circuit est en activité afin d'augmenter au maximum le courant  $I_{ds}$  et donc les performances, et au plus haut la valeur de  $V_{th}$ , afin de réduire les courants de fuite lorsque le circuit est en veille. Afin de réduire  $V_{th}$ , il suffit de polariser en direct le body (FBB pour « Forward Body Biasing ») avec  $V_{bs} > 0$  pour les NMOS et  $V_{bs} < 0$  pour les PMOS ; à l'opposé, afin d'augmenter  $V_{th}$ , il suffit de polariser en inverse (RBB pour « Reverse Body Biasing ») avec  $V_{bs} < 0$  pour les NMOS et  $V_{bs} > 0$  pour les PMOS. La figure montre trois polarisations différentes d'un MOS de type N et leur impact sur le courant de canal  $I_{ds}$  en fonction de la différence de potentiel  $V_{gs}$ .



**figure : Impact sur  $I_{ds}$  de la polarisation de body pour le cas d'un transistor NMOS LP SVt à 125°C**

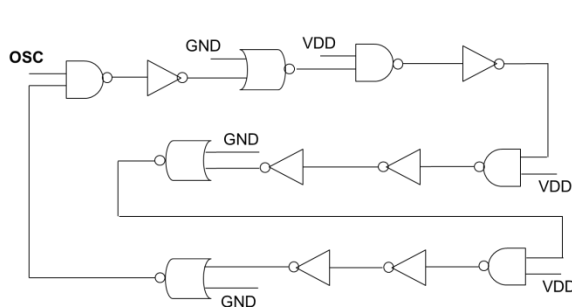


La polarisation en direct se limite à la tension de coude de la diode,  $V_D$ , de jonction entre le body et la source. Il existe deux manières de polariser le body, comme le souligne la figure : soit en modifiant la polarisation directement depuis l'extérieur du circuit via une seule alimentation (VDDS dans le cas d'un transistor PMOS), soit en polarisant d'une manière dynamique, contrôlée extérieurement ou de manière automatique via une unité de gestion de puissance [Mou'10].



**figure : Polarisation du body temporelle ou dynamique pour le cas d'un transistor PMOS**

La fréquence d'oscillation ainsi que la puissance statique et dynamique d'un oscillateur en anneau de technologie 65nm BULK et de type HVt sont présentées figure . Une polarisation des substrats des caissons des NMOS et PMOS, FBB et RBB de 150mV est comparée dans le cas où aucune polarisation n'est appliquée  $|V_{bs}|=0$ .



**Simulation à 25°C ; Polarisation  $\pm 150\text{mV}$**

Polarisation	Puissance (W)		Délai
	statique	dynamique	
RBB	1,22e-10	4,17e-5	893,5ps
$ V_{bs} =0$	1,32e-10	4,64e-5	786,3ps
FBB	1,74e-10	5,08e-5	714,1ps

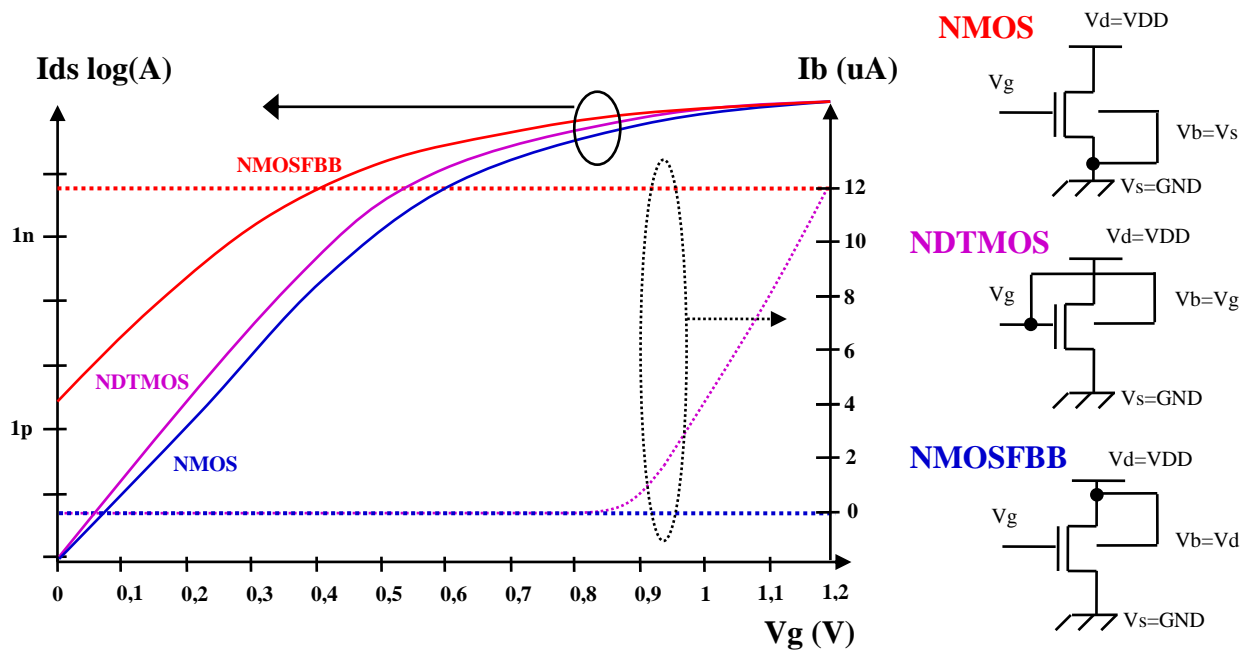
**figure : Oscillateur en anneau, délai et consommation en fonction de la polarisation de substrat en BULK 65nm LP HVt**

Le gain en fréquence en mode FBB est de 10% par rapport à une polarisation à 0 et de 20% par rapport au mode RBB. L'inconvénient est une augmentation des courants de fuite respective de 30% et 43%. Le produit de la puissance dynamique, lorsque  $\text{OSC}=1$ , et du délai reste sensiblement le même dans les trois cas de figure.

En BULK, la prise body est généralement une prise caisson commune, cette technique est donc facilement réalisable pour n'importe quel circuit de cette technologie. En PD-SOI, cette technique n'est pas applicable pour les transistors de type FB, puisqu'ils n'ont pas de body accessible. En revanche, si les transistors constituant la partie logique sont à prise body, cette technique est envisageable mais beaucoup trop contraignante. En effet, la prise body en PD-SOI introduit une surface supplémentaire par rapport au BULK, entraînant une augmentation de la capacité de grille et donc une perte en performances ; de plus, l'accès individuel de chaque prise body des transistors induit un routage contraignant et supplémentaire.

#### 2.3.2.6. Polarisation Dynamique du body

Basée sur la technique VTCMOS, la technique appelée Dynamic Threshold MOS « DTMOS » [Ass'94], pour MOS à tension de seuil dynamique en français, utilise le potentiel de grille comme potentiel de polarisation de substrat. Ainsi le potentiel de substrat suit le potentiel de la grille et évolue en fonction du mode de fonctionnement du transistor. Lorsque le transistor, NMOS ou PMOS, est bloqué,  $|V_{gs}|=|V_{bs}|=0$  la tension de seuil  $|V_{th}|$  devient élevée, induisant un courant de fuite faible. A l'opposé, lorsque le transistor est passant  $|V_{gs}|=|V_{bs}|=V_{DD}$ , le potentiel de substrat est fortement polarisé en direct, « FBB ». Dans ce cas, la tension de seuil  $|V_{th}|$  est réduite : cela engendre un courant  $I_{ds}$  et des performances plus élevées. En comparaison avec une polarisation statique nulle,  $|V_{bs}|=0$ , la pente sous seuil d'un transistor monté en DTMOS est bien meilleure ( $\approx 75\text{mV/décade}$ ), comme le montre la figure .



**figure : Courant de drain et de body pour 3 configurations de polarisation. Transistors de dimension  $W=300nm$   $L=60nm$  en technologie PD-SOI 65 LP HVt pour  $V_{DD}=1,2V$  à  $25^{\circ}C$ .**

Lorsque le transistor « NDTMOS » est bloqué,  $|V_{gs}|=0$ , le body est polarisé à  $|V_{bs}|=0$  de la même manière que le transistor « NMOS » l'est ; ils ont donc les mêmes courants de fuite  $I_{ds}$  et  $I_b$ . A l'opposé lorsque le transistor « NDTMOS » est passant, le body est polarisé de la même manière que le MOS « NMOSFBB » ; ils ont donc les mêmes courants  $I_{ds}$  et  $I_b$ . L'inconvénient dans le second cas de figure est le courant de body,  $I_b$ , qui devient important voir supérieur au courant  $I_{ds}$ . Ce courant est le courant direct de diode entre le body et la source. C'est pour cela que la technique DTMOS implique une tension d'alimentation  $V_{DD}$  inférieure à  $V_D$  (généralement  $< 0,6V$ ) afin d'éviter que la diode entre le body et la source soit passante.

En BULK, cette technique nécessite la mise en place d'une isolation de chacun des caissons des transistors, ce qui est compliqué et trop gourmand en surface. En PD-SOI, étant donné que tous les transistors à prise substrat ont leurs « body » isolés les uns des autres, cette solution s'applique assez aisément. Bien sûr les transistors BC nécessitent une surface d'implémentation supérieure aux transistors FB et ont des performances dégradées lorsque leur body est polarisé à la source ( $V_{bs}=0$ ), mais avec la technique DTMOS la polarisation dynamique permet d'augmenter la pente sous le seuil et donc les performances. Il faut aussi tenir compte de la capacité vue par la grille qui augmente en raison de la capacité de la prise de body qui s'ajoute (voir chapitre 1 : vue lay-out du DTMOS). Un facteur de mérite, présenté dans l'équation , permet de savoir si l'augmentation du courant, du DTMOS par rapport au

MOS FB, est supérieure à l'augmentation de la capacité de grille [Lho'05]. Il permet de faire le choix entre une logique FB et DTMOS. En effet si un gain existe en termes de courant de commutation  $I_{on}$ , celui-ci peut être pénalisé par une hausse de la capacité de grille  $C_{in}$  signifiant une diminution des performances et une augmentation de la consommation dynamique.

$$\frac{I_{on}(DT)}{I_{on}(FB)} \bigg/ \frac{C_{in}(DT)}{C_{in}(FB)}$$

**équation**

Le choix se portera sur une logique de type « DTMOS » si le résultat de l'équation est supérieur à 1. Et même dans ce cas, l'augmentation de la consommation dynamique et la perte en surface pousse souvent à complètement éviter cette logique pour n'utiliser que des MOS de type FB.

Le principal inconvénient de la technique « DTMOS » est la polarisation en direct de la diode Source/Substrat lorsque le transistor est actif. Le courant utile à la commutation étant le courant de drain  $I_{ds}$ , le courant  $I_b$  est un courant de fuite qui engendre une consommation totale critique pour des tensions  $V_{DD} \gg V_D$ . Cette technique est donc à utiliser pour des tensions d'alimentation inférieures ou aux alentours de  $V_D$  ( $\approx 600\text{mV}$ ).

Il existe un moyen d'éviter cette surconsommation pour des tensions  $V_{DD}$  supérieures à  $V_D$  en ajoutant au niveau du body un limiteur de courant [Ass'97] qui empêche la polarisation en directe de la diode. Un limiteur peut être implémenté de différentes manières [Cas'00], chaque solution apportant son lot d'avantages et inconvénients. La figure montre différentes configurations appuyées par des résultats de simulation du courant  $I_d$  en fonction de  $V_g$ .

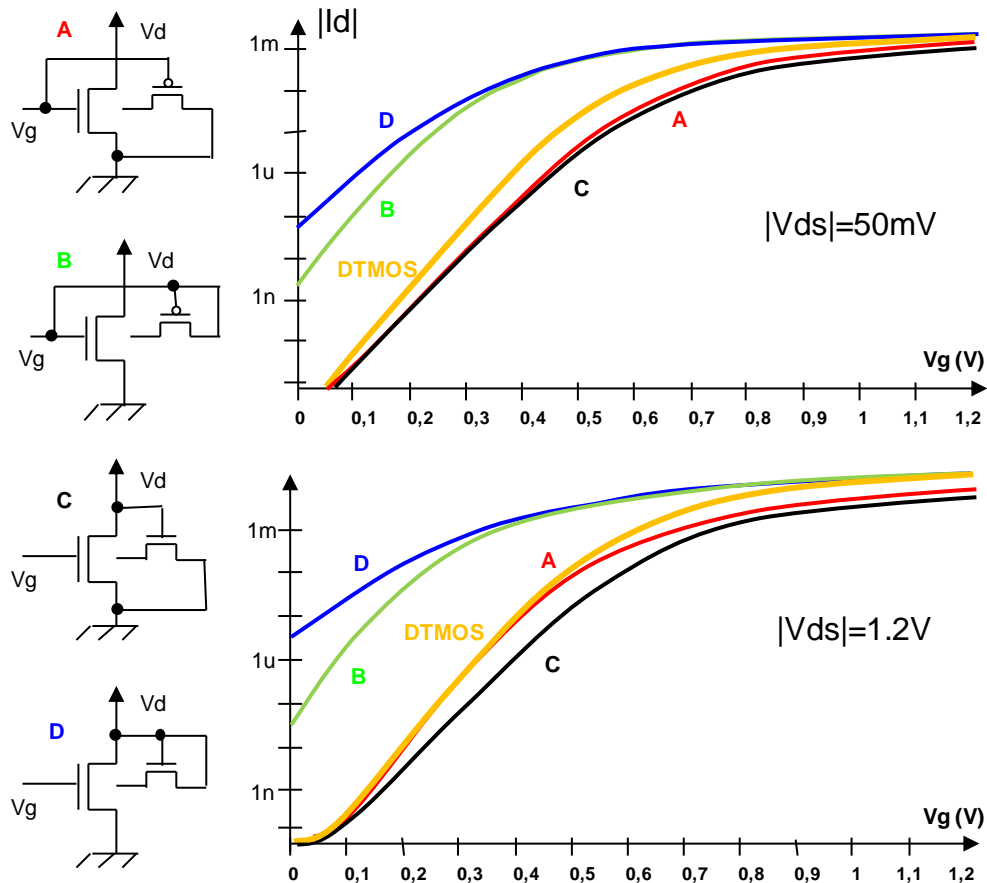


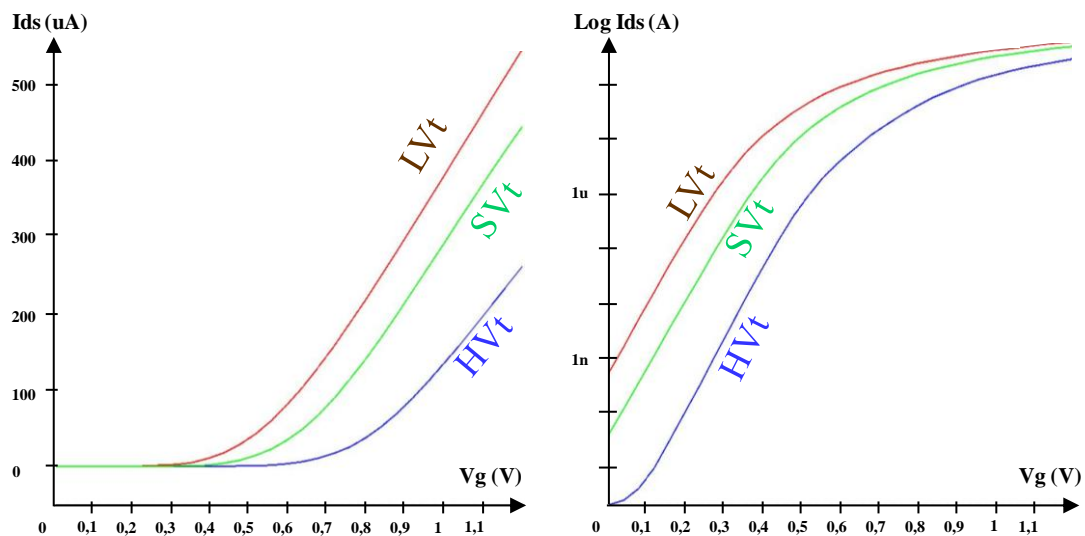
figure : DTMOS avec limiteur en PD-SOI 65nm LP HVt à 25°C

Il est évident que l'implémentation d'un limiteur pour chaque transistor induit une perte énorme en termes de surface. Il faut donc éviter d'utiliser cette solution dans la partie logique ou dans n'importe quel bloc incluant un très grand nombre de transistors, mais la réserver au contraire pour des solutions spécifiques et incluant peu de transistors. Le choix du limiteur se fera en fonction du compromis entre faible  $I_d$  à  $V_{gs}=0$ , fort  $I_d$  à  $V_{gs}=1,2\text{V}$  et faible  $I_b$  dans les deux cas. Il est préférable d'utiliser cette technique pour des transistors BC ayant une largeur très élevée car la perte en surface du limiteur devient relativement plus faible par rapport au MOS auquel il est connecté.

### 2.3.2.7. Conception Multi-Vt

Il existe un moyen d'allier hautes performances et faible consommation en co-intégrant des transistors ayant différents type de  $V_{th}$  : c'est le Multi-Vt. La tension de seuil d'un transistor  $V_{th}$  lui confère des caractéristiques électriques bien particulières. Les transistors de type « LVt » pour Low-Vt ou bas Vt auront des performances plus élevées que

leurs homologues de type « HVt » pour High-Vt ou Fort Vt, en raison des performances liées au courant  $I_{ds}$ , fonction de  $(V_{gs}-V_{th})^2$ . A contrario, les transistors de type HVt seront bien moins fuyants que les transistors de type LVt, comme le montre l'équation . Il existe un type de transistor intermédiaire, le SVt pour standard Vth qui a un Vth compris entre celui du LVt et HVt et présente donc des caractéristiques en termes de performances, consommation dynamique et statique intermédiaires. La figure montre le courant  $I_{ds}$  de 3 MOS de technologie « BULK » ayant les mêmes dimensions, à 25°C et une alimentation  $V_{ds}=V_{DD}$  pour une échelle linéaire et logarithmique.



**figure : Courant de drain  $I_{ds}$  de 3 NMOS HVt, SVt & LVt sur une échelle linéaire et logarithmique.**

La technique Multi-Vt est une technique mixant des portes logiques de différents Vth. Le principe selon [Luo'08] est le suivant : la conception d'un circuit est réalisée avec des portes logiques de type SVt. Les portes logiques des chemins critiques sont remplacées par des portes logiques de type LVt afin de minimiser les délais. Les portes logiques des chemins non critiques sont remplacées par des portes logiques de type HVt afin de réduire la consommation statique. Cette approche est similaire au mélange de portes ayant des longueurs de grille  $L_g$  différentes. [Luo'08] présente un gain de 63,8% sur la consommation statique d'un circuit microprocesseur.

## 2.4. Techniques basées sur des interrupteurs de puissance

## 2.4.1. Principe de base

La technique MTCMOS pour « Multi Threshold CMOS » est aujourd'hui la plus performante en termes de réduction de la consommation statique. Le principe est le suivant [Mut'95] : un réseau de transistors de puissance à faible consommation statique alimente une partie logique rapide et de faible consommation dynamique (figure ). Cette technique permet, par exemple, de diviser par 600 la consommation statique en technologie  $0,5\mu\text{m}$  [Mut'95].

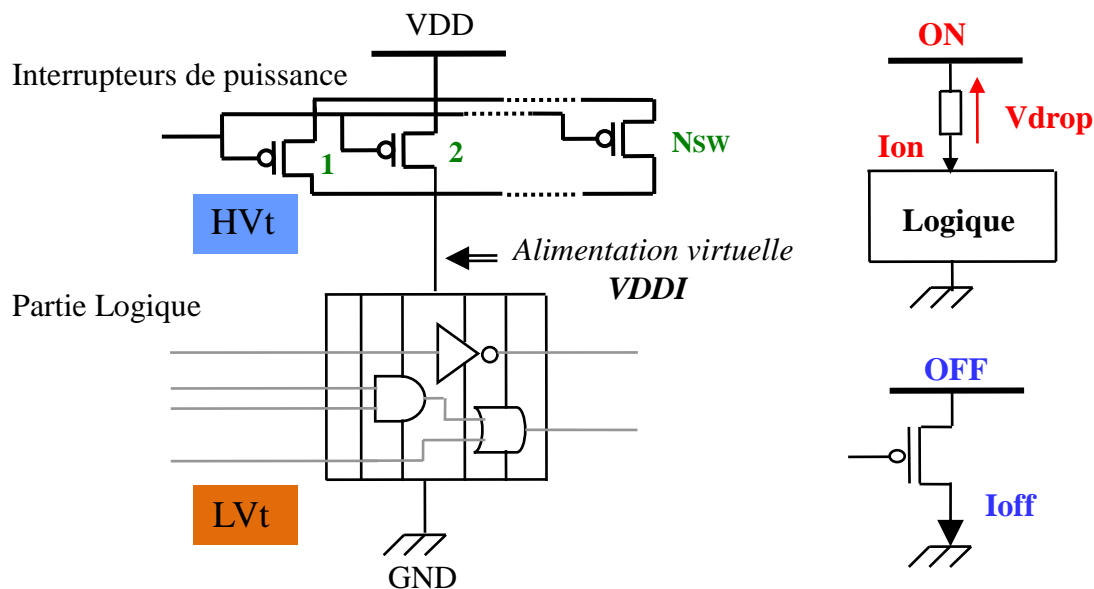


figure : Technique MTCMOS.

Le réseau d'interrupteurs de puissance (« power switch network » en anglais) peut être de type En-tête ou « Header », lorsqu'il est composé de transistors PMOS et placé entre l'alimentation VDD et l'alimentation virtuelle « VDDI », ou de type « Footer » avec des NMOS entre la masse extérieure GND et la masse virtuelle. Un circuit utilisera l'un ou l'autre mais rarement les deux simultanément. Dans le cas du « Footer » les transistors NMOS de puissance sont plus fuyants et ont une fiabilité plus faible que les transistors PMOS de la solution « Header ». C'est pour cela que l'implémentation d'un réseau de type En-tête est souvent préférée à une implémentation de type « Footer ».

Pour un réseau d'interrupteurs de puissance En-tête, lorsque celui est activé (mode « ON »), l'alimentation virtuelle de la partie logique VDDI est tirée vers VDD et le cœur logique est alors alimenté et opérationnel. Dans ce cas, la partie logique est généralement

active, et il est préférable d'utiliser une technologie de type LVt pour ses bonnes performances. Lorsque le circuit n'est pas actif, il est conseillé de désactiver le réseau d'interrupteurs de puissance (mode « OFF ») car une logique de type LVt est fortement fuyante. Dès lors, le réseau d'interrupteurs de puissance prend le relai sur la consommation statique en désactivant le cœur logique. Afin de minimiser la consommation statique du réseau d'interrupteurs de puissance, il est recommandé que celui-ci soit de type HVt car peu fuyant. C'est pour cela que cette technique s'appelle « Multi Threshold CMOS » car elle mixe le LVt d'une partie logique avec le HVt des interrupteurs de puissance. En PD-SOI, les interrupteurs peuvent être de type FB ou BC, sachant que les MOS BC sont moins fuyants. Grâce à cette technique, le gain en termes de réduction de la consommation statique est de plusieurs décades. Les inconvénients de cette technique sont la surface supplémentaire nécessaire à l'implémentation des interrupteurs de puissance et le contrôle de ceux-ci. Une légère réduction des performances liée à l'abaissement du potentiel de l'alimentation virtuelle peut apparaître lorsque le réseau d'interrupteurs de puissance est activé. Cet abaissement de l'alimentation virtuelle par rapport à l'alimentation extérieure VDD est dû à la chute de tension aux bornes de la résistance équivalente ( $R_{on}$ ) du réseau d'interrupteurs de puissance appelée « Vdrop ».

#### 2.4.2. Dimensionnement des interrupteurs

Comme son nom l'indique, un réseau d'interrupteurs de puissance est composé d'un grand nombre d'interrupteurs en parallèle, « NSW ». Le nombre d'interrupteurs unitaires, NSW, qui composent ce réseau permet de déterminer les valeurs des résistances totales  $R_{on}$  et  $R_{off}$ , et donc la chute de tension, Vdrop, et le courant de fuite  $I_{off}$  (équation ).

$$\begin{aligned} R_{totale} &= R_{unitaire} \div NSW \\ Vdrop &= R_{on_{totale}} \times I_{on} \\ I_{off} &= VDD \div R_{off_{totale}} \end{aligned}$$

**équation**

$R_{on_{unitaire}}$  est la valeur de la résistance d'un interrupteur unitaire lorsqu'il est passant. Il définit en fonction du nombre d'interrupteurs de puissance, NSW, la valeur de la résistance totale du réseau d'interrupteurs de puissance  $R_{on_{totale}}$  et donc la chute de tension « Vdrop » en



fonction du courant qui le traverse,  $I_{on}$ . A l'opposé, toujours selon NSW, lorsqu'il bloqué,  $R_{off\_unitaire}$  définit la résistance totale du réseau  $R_{off\_totale}$  et par conséquent le courant de fuite  $I_{off}$ .

Le dimensionnement consiste principalement à déterminer la longueur de grille optimale «  $L_{opt}$  » et le nombre d'interrupteurs de puissance «  $NSW_{opt}$  » donnant le meilleur compromis entre une faible résistance  $R_{on}$  et une forte résistance  $R_{off}$ . Il existe deux approches, la première considérant la chute de tension  $V_{drop}$  et le courant de fuite  $I_{off}$  [Hwa'06], la seconde la dégradation de délai [Kao'97] introduit par l'implémentation de cette technique.

### 2.4.3. Contrôle et réveil

Il existe deux modes de fonctionnement : éteint (OFF) et allumé (ON). Le contrôle de ces modes ainsi que des phases de transitions (réveil et mise en veille) est géré par un contrôleur (figure ). Ce contrôleur pilote la valeur de tension de grille des interrupteurs de puissance,  $SWVDD$ , ainsi que le potentiel de l'alimentation virtuelle  $VDDI$  [Liu'08]. Le contrôleur a aussi pour rôle de limiter l'appel en courant de la phase de réveil, évitant ainsi un pic en consommation dynamique.

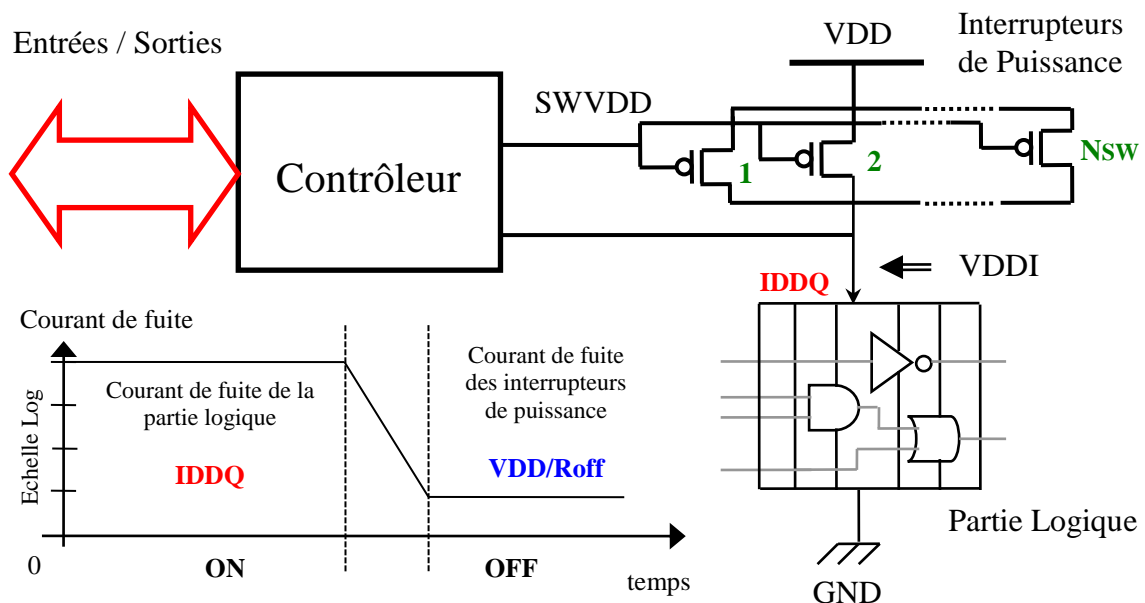


figure : Technique MTCMOS : contrôle et transition

#### 2.4.4. Implémentation

L'implémentation des interrupteurs de puissance peut être en forme d'anneau autour de la partie logique à alimenter ou répartie dans la partie logique [Lon'04]. Les avantages de la première sur la seconde sont la simplicité de l'implémentation et la facilité de la distribution de l'alimentation virtuelle. La seconde apporte une solution plus locale réduisant également la chute de tension liée à la résistance des métaux de routage ou interconnexions.

#### 2.4.5. Les Bascules de rétention

L'utilisation d'interrupteurs de puissance permettant d'alimenter ou d'éteindre une partie logique provoque la perte des données. Une solution à ce problème est de remplacer les bascules d'entrées et de sorties critiques par des bascules de rétention [Hen'05], commandées par le signal de Veille du contrôleur de la bibliothèque MTCMOS (figure ).

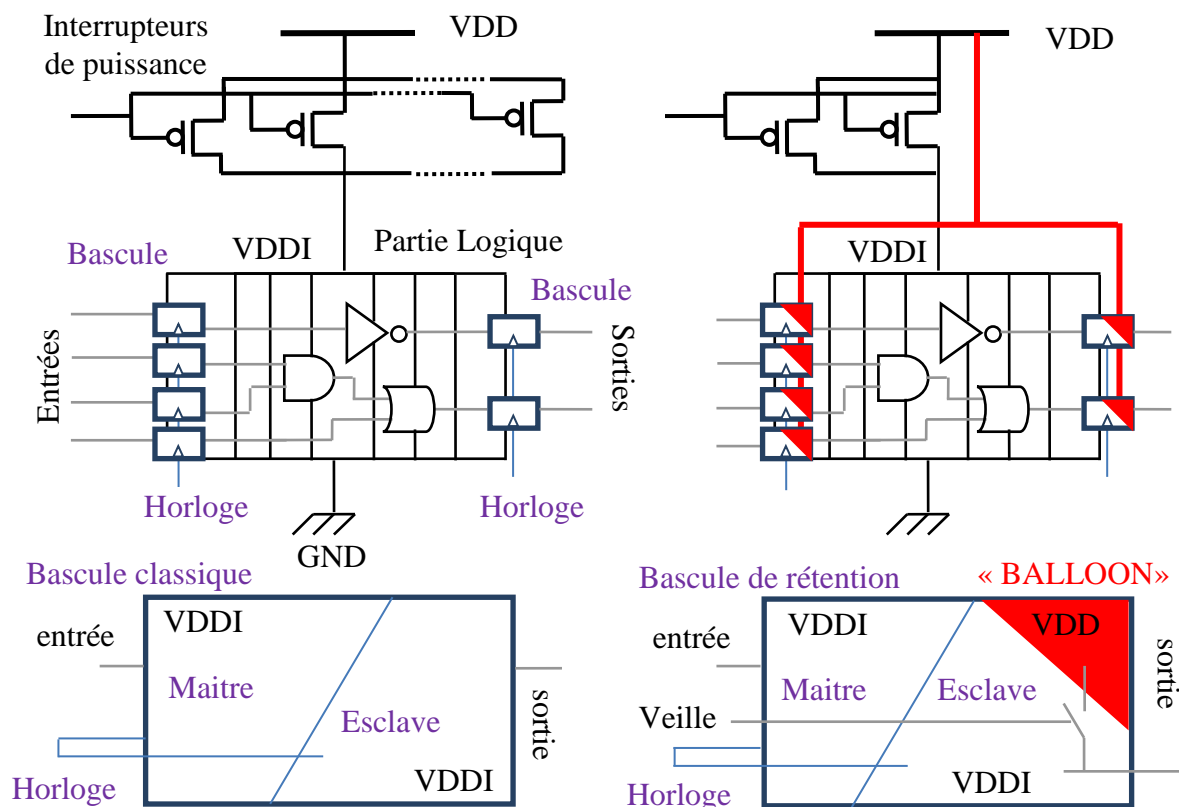


figure : Bascule de rétention associée à la technique MTCMOS

Cette bascule de rétention est divisée en deux parties, le maître et l'esclave. La particularité de cette bascule est la sous partie appelé « BALLOON » qui permet d'assurer la

rétenion car elle est toujours alimentée à VDD même quand les interrupteurs de puissance sont coupés. Lors du réveil, les données stockées dans les « BALLOON » sont redistribuées à la sortie des registres et donc à la partie logique : il n'y a pas de perte d'information et cela évite une réinitialisation.

## 2.5. Conclusion

En technologie PD-SOI, la solution qui permettrait de réduire la consommation statique de plusieurs décades semble être la technique MTCMOS. En 65nm PD-SOI l'implémentation de cette technique se ferait en utilisant de préférence des interrupteurs de puissance de type BC et HVt, peu fuyants et similaires au BULK, et en les associant à une logique de type FB, plus rapide qu'en BULK. L'optimisation de ces interrupteurs de puissance en technologie LP (Low Power) 65nm PD-SOI est détaillée dans le chapitre suivant. Une nouvelle solution d'interrupteur de puissance est proposée, appuyée par des mesures sur silicium.

- [Aga'05] Agarwal, A.; Mukhopadhyay, S.; Kim, C.; Raychowdhury, A. & Roy, K. "Leakage power analysis and reduction: models, estimation and tools" IEEE Proceedings of Computers and Digital Techniques, vol.152, p.353-368, 2005
- [Ass'94] Assaderaghi, F.; Parke, S.; Sinitsky, D.; Bokor, J.; Ko, P. & Hu, C. "A dynamic threshold voltage MOSFET (DTMOS) for very low voltage operation" IEEE Electron Device Letters, vol.15, p.510-512, 1994
- [Ass'97] Assaderaghi, F.; Sinitsky, D.; Parke, S.; Bokor, J.; Ko, P. & Hu, C. "Dynamic threshold-voltage MOSFET (DTMOS) for ultra-low voltage" IEEE Transactions on VLSI Electron Devices, vol.44, p.414-422, 1997
- [Cal'04] Calhoun, B. & Chandrakasan, A. "Characterizing and Modeling Minimum Energy Operation for Subthreshold Circuits" International Symposium on Low Power Electronics and Design, p.90-95, 2004
- [Cas'00] Casu, M.; Masera, G.; Piccinini, G.; Ruo Roch, M. & Zamboni, M. "Comparative analysis of PD-SOI active body-biasing circuits" IEEE SOI Conference, p.94-95, 2000
- [Cha'92] Chandrakasan, A.; Sheng, S. & Brodersen, R. "Low-power CMOS digital design" IEEE Journal of Solid-State Circuits, vol.27, p.473-484, 1992
- [Cla'04] Clark, L.; Patel, R. & Beatty, T "Managing Standby and Active Mode Leakage Power Sub-micron Design" International Symposium on Low Power Electronics and Design, p.274-279, 2004
- [Hen'05] Henzler, S.; Nirschi, T.; Pacha, C.; Spindler, P. et al. "Dynamic state-retention flip-flop for fine-grained sleep-transistor scheme" European Solid-State Circuits Conference, p.145-148, 2005
- [Hor'94] Horowitz, M.; Indermaur, T. & Gonzalez, R. "Low-power digital design" IEEE Symposium on Low Power Electronics, p.8-11, 1994
- [Hwa'06] Hwang, C.; Kang, C. & Pedram, M "Gate sizing and replication to minimize the effects of virtual ground parasitic resistances in MTCMOS designs" Quality Electronic Design, 2006
- [Im'03] Im, H.; Inukai, T.; Gomyo, H.; Hiramoto, T. & Sakurai, T. "VTCMOS characteristics and its optimum conditions predicted by a compact analytical model" IEEE Transactions on Very Large Scale Integration Systems, vol.11, p.755-761, 2003
- [Joh'02] Johnson, M.; Somasekhar, D.; Chiou, L.Y. & Roy, K. "Leakage control with efficient use of transistor stacks in single threshold CMOS" IEEE Transactions on Very Large Scale Integration Systems, vol.10, p.1-5, 2002
- [Kao'97] Kao, J.; Chandrakasan, A. & Antoniadis, D. "Transistor Sizing Issues And Tool For Multi-threshold Cmos Technology" Design Automation Conference, 1997
- [Kea'07] Keating, M. et al. "Low Power Methodology Manual For System-on-Chip Design" Springer, 2007

- [Lho'05] L'Hostis, N.; Valentian, A. & Amara, A. "A 130nm partially depleted SOI technology menu for low-power applications" IEEE NEWCAS Conference, p.175-178, 2005
- [Lin'09] Lin, B.; Mallik, A.; Dinda, P.; Memik, G. & Dick, R. "User- and process-driven dynamic voltage and frequency scaling" IEEE International Symposium on Performance Analysis of Systems and Software, p.11-22, 2009
- [Liu'08] Liu, Z. & Kursun, V. "Characterisation of wake-up delay versus mode power consumption and sleep/active mode transistion energy overhead tradeoffs in MTCMOS circuits" Symposium on Circuits and Systems, NWSCAS, p.362-365, 2008
- [Lon'04] Long, C. & He, L. "Distributed sleep transistor network for power reduction" IEEE transactions on Very Large Scale Integration Systems, vol.12, p.937-946
- [Luo'08] Luo, T.; Newmark, D. & Pan, D. "Total power optimization combining placement, sizing and multi-Vt through slack distribution management" Asia and South Pacific Design Automation Conference, p.352 -357, 2008
- [Mah'09] Mahmoodi, H.; Tirumalashetty, V.; Cooke, M. & Roy, K. "Ultra Low-Power Clocking Scheme Using Energy Recovery and Clock Gating" IEEE Transactions on Very Large Scale Integration, vol.17, p.33-44, 2009
- [Mou'10] Moubdi, N.; Maurine, P.; Wilson, R.; Azemard, N.; Engels, S.; Rolindez, L. & Heinrich, V. "Voltage Scaling and Body Biasing Methodology for High Performance Hardwired LDPC" IEEE International Conference on ICICDT, P.82-85, 2010
- [Mut'95] Mutoh, S.; Douseki, T.; Matsuya, Y.; Aoki, T.; Shigematsu, S. & Yamada, J. "1-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS" IEEE Journal of Solid-State Circuits, vol.30, p.847-854, 1995
- [Nar'01] Narendra, S.; Borkar, S.; De, V.; Antoniadis, D. & Chandrakasan, A. "Scaling of stack effect and its application for leakage reduction" International Symposium on Low Power Electronics and Design, p.195-200, 2001
- [Pig'04] Piguet, C.; Schuster, C.; Nagel, J.-L. "Optimizing architecture activity and logic depth for static and dynamic power reduction" NEWCAS , p. 41-44, 2004
- [Suz'00] Suzuki, K. "Short channel MOSFET model using a universal channel depletion width parameter" IEEE Transactions on Electron Devices, vol.47, p.1202-1208, 2000
- [Van'04] Van der Meer, P.R.; Van Staveren, A. & Van Roermund, A.H.M "Low-Power Deep Sub-Micron CMOS Logic" Kluwer Academic, 2004





### 3.Optimisation des interrupteurs de puissance en 65nm PD-SOI

*Le changement de contexte obligeant les concepteurs à utiliser des interrupteurs de puissance en technologie PD-SOI 65nm est introduit dans ce chapitre. L'optimisation selon l'état de l'art, la méthodologie de dimensionnement ainsi que l'implémentation de ces interrupteurs sont présentées et appuyées par des résultats de simulation. Un nouveau facteur de mérite, prenant en compte des caractéristiques électriques et de la surface y est proposé pour la sélection du meilleur interrupteur de puissance. Une solution innovante d'interrupteur de puissance, basée sur un bloc de polarisation dynamique et auto adaptatif, est présentée et détaillée. Ce montage apporte pour le même courant de fuite  $I_{off}$ , en mode éteint, une réduction de la résistance équivalente en mode passant,  $R_{on}$ , de 20%. La bibliothèque associée à ce nouveau montage a été validée à travers un circuit de test, ses paramètres électriques mesurés et comparés à une solution de référence ainsi qu'au circuit correspondant en technologie BULK.*



### 3.1. Evolution des contraintes technologiques et problèmes spécifiques liés au PD-SOI

La technologie PD-SOI a plus de fuites que la technologie BULK en raison de l'effet du body flottant de ses transistors FB. Ceci est vrai lorsque la tension de seuil  $V_{th}$  en régime linéaire est sensiblement la même dans les deux technologies.

#### 3.1.1. Besoin d'interrupteurs de puissance en 65nm

L'évolution des technologies montre que la consommation statique devient non-négligeable, voir égale ou supérieure à la consommation dynamique selon les types de circuits et d'applications. Alors qu'en 90nm le ratio entre la consommation dynamique et la consommation statique est aux alentours de l'unité, à partir du 65nm la consommation devient nettement supérieure à la consommation dynamique dans certains cas de figure où aucune technique de réduction n'est implémentée. Le tableau montre l'augmentation, prédite par l'ITRS et reportée dans [Kea'07], de la consommation dynamique, statique et totale par unité de surface de circuits de type SOC (System On Chip), de la technologie 90nm à la technologie 45nm.

Technologie	90nm	65nm	45nm
Consommation dynamique / $m^2$	×1	×1,4	×2
Consommation statique / $m^2$	×1	×2,5	×6,5
Consommation totale / $m^2$	×1	×2	×4

**tableau : Evolution de la consommation par unité de surface normalisée des technologies 90nm vers 65nm et 45nm**

La consommation statique augmente de manière exponentielle à la différence de la consommation dynamique qui croît de manière linéaire. Les prédictions de l'ITRS et les différents résultats présentés au cours de cette dernière décennie sur cette augmentation dramatique des courants de fuite (figure [Qi'06]), ont obligé les concepteurs à introduire des techniques de réduction de la consommation statique. Que ce soit en technologie PD-SOI ou BULK, la technique MTCMOS est devenue indispensable, en particulier, pour les circuits dédiés aux applications faible consommation.

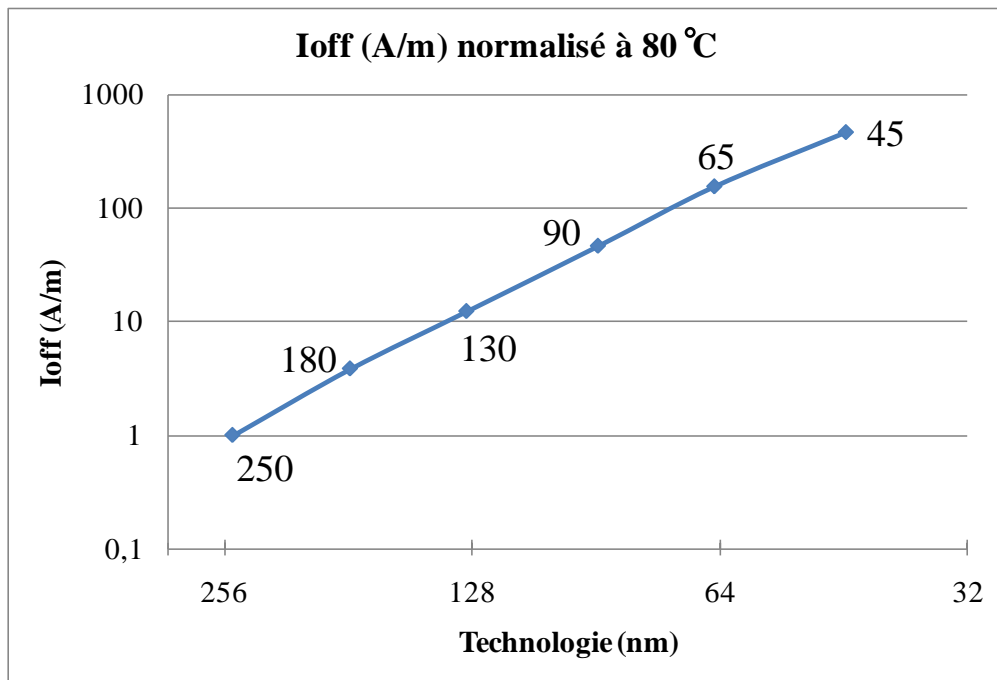


figure : Evolution du courant de fuite Ioff normalisé avec les technologies à 80°C

### 3.1.2. Contexte de la technologie 0,13µm PD-SOI

En 0,13µm PD-SOI, les différentes tensions de seuils proposées, LVt (faible Vth), HVt (haut Vth) et SVt (standard Vth), ne sont pas les mêmes que celles proposées en technologie BULK. En effet, l'objectif n'est pas d'aligner les tensions de seuil entre elles mais plutôt les différents courants de saturation (Idsat) ou performances des MOS (ex : délai de l'inverseur). En 0,13µm, une tension de seuil Vth plus élevée en technologie PD-SOI par rapport à la technologie BULK va permettre d'atteindre les mêmes performances tout en réduisant la consommation statique [Lho'05]. D'une manière similaire, en réduisant légèrement cette même tension de seuil en valeur absolue afin de viser les mêmes courants de fuite que la technologie BULK, la technologie PD-SOI devient plus performante. C'est ce deuxième choix qui a été préféré chez STMicroelectronics.

### 3.1.3. Comparaison des technologies PD-SOI et BULK en 65nm

[Fla'09] a montré que le passage d'une technologie BULK vers du PD-SOI en 65nm ou 90nm permet de gagner 20% sur le délai global d'un circuit pour une consommation dynamique quasi similaire, mais au prix d'une augmentation croissante de la consommation statique par rapport au BULK (tableau ).

<b>SOI vs BULK @ 1.2V @ 125 °C</b>	<b>90nm</b>	<b>65nm</b>
Délai (ps)	<b>- 18%</b>	<b>- 20%</b>
Puissance dynamique ( $\mu$ W)	<b>- 4%</b>	<b>=</b>
Puissance statique (nW)	<b><math>\times 2,3</math></b>	<b><math>\times 5</math></b>

**tableau : Comparaison du délai et de la consommation entre les technologies PD-SOI et BULK d'un oscillateur en anneaux en 65nm et 90nm [Fla'09]**

Pour les technologies submicroniques, et plus particulièrement en 65nm, aligner les tensions de seuil pour obtenir les mêmes performances ou, comme cela a été en 0,13 $\mu$ m chez STMicroelectronics, les mêmes courants de fuite, ne permet pas à la technologie PD-SOI d'être, en même temps, moins fuyante et plus performante.

En technologie faible consommation 65nm PD-SOI, la tension de seuil, sans polarisation de « body » en régime linéaire, est légèrement plus élevée qu'en technologie BULK. Cependant, comme cette tension de seuil est déjà faible, (en raison de la tension d'alimentation de ce nœud technologique), l'effet de body flottant qui l'abaisse a un impact important sur les courants de fuites. Cette tension de seuil devient plus faible que celle de la technologie BULK en régime saturé, entraînant un courant de fuite plus élevé qui augmente de manière exponentielle avec la réduction des nœuds technologiques.

Il faut donc considérer que la technologie 65nm PD-SOI ne peut pas gagner sur tous les tableaux par rapport à la technologie BULK en étant à la fois plus rapide tout en ayant des consommations dynamiques et statiques inférieures.

Le choix retenu a été d'amener les circuits à fonctionner plus vite, ou à consommer moins dynamiquement à vitesse constante. Ce choix induit une consommation statique supérieure à celle du BULK. Heureusement des techniques permettant de réduire la consommation statique, notamment la technique MTCMOS associant à la partie logique du circuit des interrupteurs de puissance basés sur des MOS peu fuyants, existent et sont aussi mises en œuvre en technologie CMOS BULK. C'est l'association de ces techniques de conception et des avantages de la technologie PD-SOI qui permet de fabriquer des circuits ayant globalement de meilleures performances et une faible consommation électrique.

### 3.1.4. Contrainte supplémentaire apportée par l'effet de body flottant sur les interrupteurs de puissance

En 65nm PD-SOI, l'effet de body flottant impacte fortement la consommation statique. Utiliser des interrupteurs de puissance de type body flottant (FB) ne permet donc pas d'obtenir un gain par rapport au BULK en terme de consommation statique, à moins d'utiliser soit une tension de seuil  $V_{th}$  plus élevée, soit une tension,  $V_{ds}$ , aux bornes du transistor plus faible (ce qui permet de réduire l'effet de body flottant). Ces deux options induiraient une dégradation des performances trop importante.

Une autre solution consiste à utiliser des interrupteurs de puissance de type body contacté (BC). Comme ceux-ci sont alignés sur les transistors en technologie BULK, ils ont le même courant de fuite et le même courant de saturation. Les interrupteurs de puissance ont un fonctionnement quasi-statique, ON et OFF. Ils sont alimentés ou éteints et ceci pendant de longues périodes par rapport au cycle d'horloge. Ce n'est donc pas leur léger excès de consommation dynamique qui apparaît lors du réveil, dû à une capacité de grille légèrement supérieure, qui pose un problème. Ils ont également une largeur de transistor élevée, ce qui fait que l'utilisation de BC n'impacte relativement pas la surface d'implémentation totale (<10%). Quant à la période de transition, elle peut être lente et progressive, afin de prévenir un appel en courant trop élevé, ou relativement plus rapide ; mais cette durée de transition reste bien inférieure à la durée des modes ON et OFF.

### 3.2. Conception et optimisation des interrupteurs de puissance, selon l'état de l'art

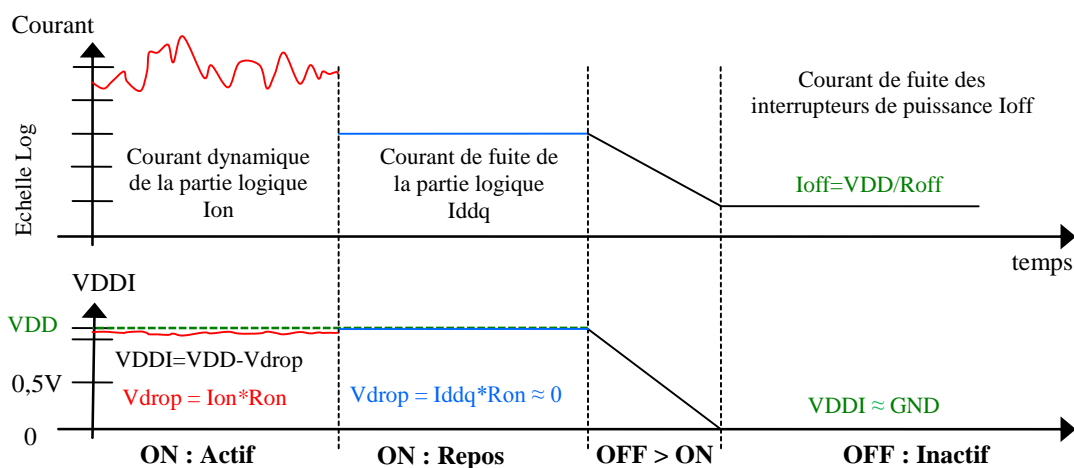
Dans un premier temps, nous avons adapté et optimisé les interrupteurs de puissance proposés dans la littérature et dans l'industrie à la technologie 65nm LP PD-SOI. Plusieurs paramètres, comme les résistances équivalentes du réseau d'interrupteurs de puissance et la surface d'implémentation, ont été améliorés. Une méthodologie de sélection des interrupteurs de puissance prenant en compte les éléments parasites, est également décrite et proposée.

#### 3.2.1. Introduction des paramètres électriques à optimiser : $R_{on}$ et $R_{off}$

Le réseau d'interrupteurs de puissance est comparable, au premier ordre, à une résistance de valeur  $R_{on}$  ou  $R_{off}$ , respectivement lorsque celui est passant ou bloqué.  $R_{on}$  et  $R_{off}$  sont inversement proportionnelles au nombre de transistors unitaires « NSW » mis en

parallèle, qui composent le réseau, ainsi qu'aux résistances équivalentes du transistor unitaire faisant office d'interrupteur de puissance,  $R_{on\_unitaire}$  et  $R_{off\_unitaire}$ . La figure illustre les trois modes de fonctionnement des interrupteurs de puissance avec leur impact sur le courant et l'alimentation virtuelle du circuit « VDDI ».

Le mode actif, « ON », est le mode où les interrupteurs de puissance sont passants. Dans ce cas, le circuit est alimenté et en fonctionnement ; l'appel en courant,  $I_{on}$ , lié à la commutation des transistors de la partie logique entraîne une chute de tension, «  $V_{drop}$  », impactant la valeur de l'alimentation virtuelle, VDDI. Cette chute de tension est fonction de la résistance équivalente du réseau d'interrupteurs de puissance,  $R_{on}$ , lorsque celui-ci est passant. Le mode repos signifie que les interrupteurs de puissance sont passants mais que le circuit ne commute pas. Le courant  $I_{ddq}$  est le courant de fuite de la partie logique en mode repos. Il est inférieur à  $I_{on}$  et induit une chute de tension,  $V_{drop}$ , nettement inférieure à la précédente. Dans le dernier mode inactif ou de veille, « OFF », les interrupteurs de puissance sont éteints afin de couper au maximum le courant de fuite de la partie logique en mode repos, passant ainsi d'un courant  $I_{ddq}$  à  $I_{off}$ . L'alimentation virtuelle, VDDI, passe alors d'une valeur proche de VDD à GND. Le courant de fuite  $I_{off}$  dépend de la résistance équivalente des interrupteurs de puissance bloqués,  $R_{off}$ , et de la tension d'alimentation VDD.



**figure : Impact du réseau d'interrupteurs de puissance sur le courant et l'alimentation du circuit en mode ON et OFF**

Un des premiers objectifs d'optimisation est de proposer une solution d'interrupteurs de puissance ayant un rapport  $R_{off}/R_{on}$  le plus élevé possible ainsi qu'une méthode pour y arriver.  $R_{off}/R_{on}$  est égal à  $R_{off\_unitaire}/R_{on\_unitaire}$ . Ce critère vise, à la fois, un courant de fuite,  $I_{off}$ , faible en mode inactif, et une chute de tension,  $V_{drop}$ , minimale en mode actif.

La chute de tension réduisant la tension d'alimentation virtuelle, VDDI, introduit une dégradation des temps de propagation par rapport au même circuit sans technique MTCMOS. Cependant, comme le courant de fuite de la partie logique peut être coupé, il est recommandé, et donc souvent utilisé, une tension de seuil très basse (LVt) dans la partie logique, ce qui permet d'atteindre des vitesses inatteignables avec une tension de seuil haute ou standard. A titre d'illustration, une comparaison de trois cas de figure est présentée tableau . Il montre une dégradation de délai, proche de +1%, induite par l'insertion d'interrupteurs de puissance dans un oscillateur en anneau LVt, pour une réduction de la consommation statique de 3 décades. Il montre également que cet oscillateur est à la fois plus rapide que son équivalent conçu en HVt sans interrupteurs de puissance et moins consommant en mode OFF.

<b>Fréquence d'un Oscillateur en Anneau 65nm LP PD-SOI</b>	<b>LVt référence</b>	<b>LVt avec "HVt power switch"</b>	<b>HVt</b>
Délai	<b>1</b>	<b>+ 1%</b>	<b>+75%</b>
Puissance dynamique	<b>1</b>	<b>- 1%</b>	<b>- 35%</b>
Puissance statique	<b>1</b>	<b>/ 1000</b>	<b>/ 50</b>

**tableau : Comparaison du délai et de la consommation de trois oscillateurs en anneaux en LP PD-SOI 65nm normalisés.**

Abaissier au maximum le courant de fuite en mode OFF implique de réduire le plus possible, le nombre d'interrupteurs de puissance, NSW. Malheureusement, ceci aura pour effet d'augmenter la résistance Ron, et donc la chute de tension, Vdrop, ce qui va dégrader la vitesse du circuit. La recherche du meilleur compromis entre faible Vdrop et faible courant de fuite Ioff, paramètres opposés, est donc nécessaire.

### 3.2.2. Optimisation des solutions proposées dans l'état de l'art

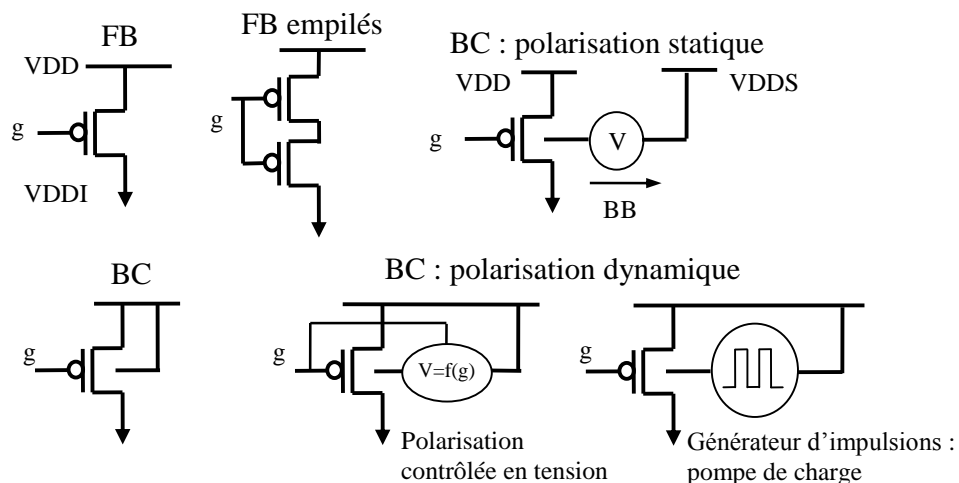
#### 3.2.2.1. Description des solutions étudiées

Il existe plusieurs solutions possibles en technologie PD-SOI. La première est d'utiliser un interrupteur de puissance de type body flottant (FB), à haute tension de seuil. Une première optimisation consiste, dans ce cas, à travailler sur la longueur de grille du transistor unitaire afin d'obtenir le ratio Roff/Ron le plus élevé. Cependant, cette solution n'est pas la meilleure en raison de l'effet de body flottant qui abaisse la tension de seuil et

réduit  $R_{off}$ . Pour pallier ce problème, il est possible d'empiler des transistors FB, ce qui permet de réduire le  $V_{ds}$  de chacun des MOS empilés et donc de minimiser cet effet flottant. Au-delà de deux MOS empilés, le ratio  $R_{off}/R_{on}$  s'écroule [Das'03] parce que  $R_{off}$  reste sensiblement le même alors que  $R_{on}$  est divisé autant de fois qu'il y a de transistors empilés.

Une seconde approche est d'utiliser un transistor à body contacté (BC), également à haute tension de seuil. De même que précédemment, il suffit d'adapter la longueur de grille pour obtenir la solution avec le meilleur ratio  $R_{off}/R_{on}$  par unité de largeur. Comme pour le nombre d'interrupteurs de puissance en parallèle,  $N_{SW}$ ,  $R_{on}$  et  $R_{off}$  sont proportionnels à la largeur de grille,  $W$ .

Il est également possible d'optimiser la meilleure solution BC, en polarisant statiquement ou dynamiquement le body des transistors, en direct (FBB) afin de réduire la résistance équivalente  $R_{on}$  et en inverse (RBB) afin d'augmenter la résistance équivalente  $R_{off}$ . Toutes ces solutions sont illustrées figure .



**figure : Présentation des différentes solutions d'interrupteur de puissance de type En-tête**

La solution basée sur la pompe de charge a un fonctionnement coûteux en termes de consommation dynamique et la mesure de sa résistance  $R_{on}$  varie temporellement [Tho'09]. Comparée à des solutions plus simples, cette solution n'est pas retenue car elle nécessite également un générateur d'impulsion, ce qui la rend plus complexe.

Pour comparer les possibilités restantes, une première étape est d'analyser les valeurs de  $R_{on}$  et de  $R_{off}$  et plus particulièrement le ratio  $R_{off}/R_{on}$ , pour plusieurs tensions d'alimentation, longueurs de grille, températures, tout en utilisant la même largeur de transistor.

### 3.2.2.2. Choix de la structure pertinente

Le premier choix est celui de la tension de seuil. Quelle est la tension de seuil  $V_{th}$  qui permet d'obtenir le meilleur ratio  $I_{on}/I_{off}$  ? Il apparaît qu'il est préférable d'utiliser une tension de seuil haute (HVT) car son ratio  $I_{on}/I_{off}$  est plus élevé par rapport à celui des tensions de seuil standard et faible, SVt et LVt.

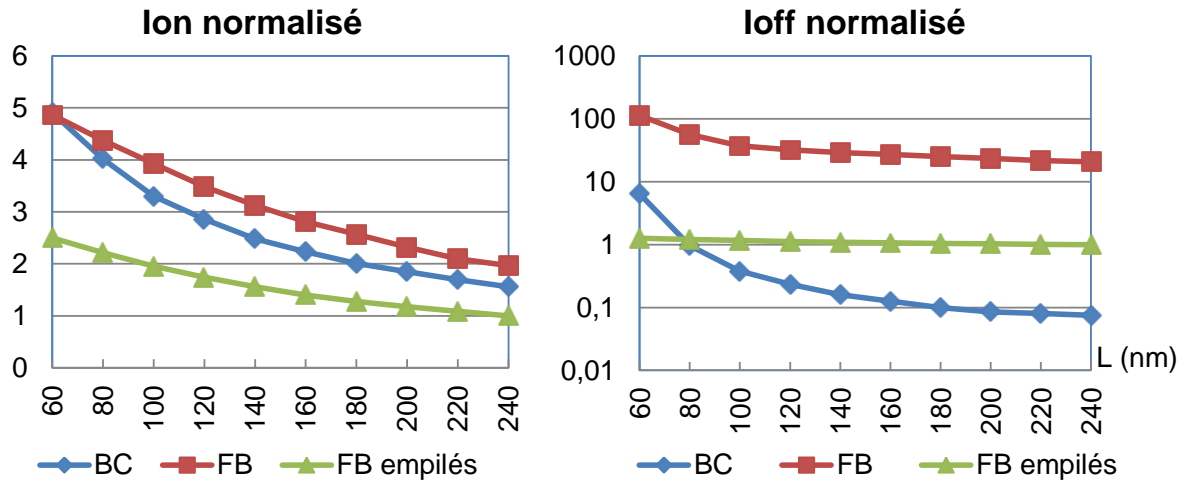
Ensuite, une seconde comparaison est faite en ne considérant que les interrupteurs de puissance ayant une tension de seuil  $V_{th}$  qui ne peut pas être modulée par l'utilisateur. Une optimisation, par la suite, inclura une polarisation de body. Pour l'instant, seuls les transistors FB et BC avec le body relié à la source sont concernés. Ces solutions, sans polarisation de body, ne nécessitent pas d'alimentation supplémentaire.

La résistance de l'interrupteur de puissance est mesurée par la relation  $V_{ds}/I_{ds}$  ( $R_{on}$  par  $V_{dson}/I_{on}$  et  $R_{off}$  par  $V_{dsoff}/I_{off}$ ).  $V_{dson}$  est la chute de tension aux bornes de l'interrupteur de puissance lorsque celui-ci est passant (mode ON), et vaut au maximum quelques dizaines de millivolts ; elle correspond à  $V_{drop}$ . L'interrupteur de puissance, lorsqu'il est passant, se doit être dans un mode linéaire, pour être le plus efficace et le moins résistif. C'est dans ce mode de fonctionnement, pour une différence de potentiel  $V_{dson}$  compris entre 0 et 50mV, qu'il est comparable à une résistance.  $I_{on}$  est le courant traversant le réseau d'interrupteurs de puissance, il représente le courant dynamique de la partie logique en pleine activité. Mais  $I_{on}$  peut être aussi considéré comme le courant que peuvent délivrer les interrupteurs de puissance pour une chute de tension fixe, calculé par  $V_{dson}/R_{on}$ .

$V_{dsoff}$  est la chute de tension lorsque le réseau d'interrupteurs de puissance est bloqué (mode OFF) ; dans ce cas  $V_{dsoff}$  vaut VDD puisque l'alimentation virtuelle, VDDI, est proche de GND (0V).  $I_{off}$  est le courant de fuite du réseau d'interrupteurs de puissance lorsqu'il est éteint ; il ne dépend pas de la partie logique.

Dans le cas de notre étude, le ratio  $R_{off}/R_{on}$  est égal à  $I_{on}/I_{off}$  si et seulement si, pour mesurer ces deux courants, les mêmes chutes de tensions sont appliquées aux cas considérés : même  $V_{dson}$  dans le mode ON et même  $V_{dsoff}$  dans le mode OFF.





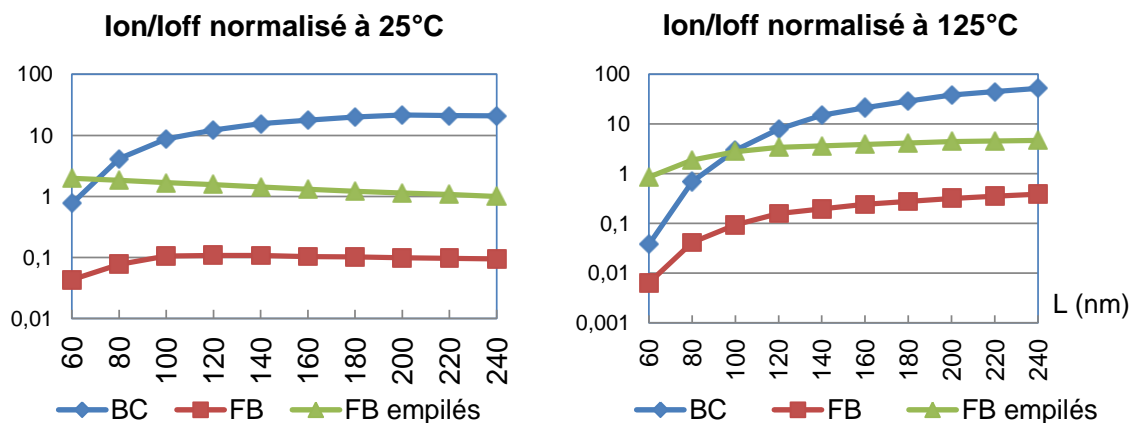
**figure : Ion et Ioff à 25°C pour un interrupteur de puissance de type BC, FB et FB empilés (2 MOS) pour une tension VDD=1,2V**

La figure illustre l'évolution des courants Ion et Ioff, pour une tension d'alimentation VDD=1.2V, une chute de tension Vdson=25mV, une seule et même largeur de grille W, et pour plusieurs longueurs de grille L. Les résultats sont normalisés vis-à-vis de l'interrupteur de puissance constitué de deux transistors FB empilés ayant une longueur de grille de L=240nm. La largeur de grille maximale du transistor BC est de W=5µm avec une prise body sur chaque extrémité. Au-delà de cette largeur, un effet de body flottant peut réapparaître à cause de la résistance d'accès élevée de la prise body.

Une comparaison du courant Ion montre que, pour la même longueur de grille L, le transistor FB permet de délivrer à la partie logique un courant Ion plus élevé par rapport aux autres solutions et pour la même chute de tension à ses bornes (Vdson=25mV). Il est suivi par le courant du transistor BC. Pour une longueur de grille L=60nm, l'effet de body flottant a moins d'impact : Vds est faible et l'effet de canal court du transistor devient plus important, dominant l'effet de body flottant. Le courant Ion de l'interrupteur de puissance à MOS FB empilés est environ deux fois plus faible que les deux autres, FB et BC. Ceci s'explique par une résistance unitaire de l'interrupteur de puissance totale, Ron, deux fois plus élevée due à la mise en série de deux MOS.

Le courant Ioff de l'interrupteur de puissance FB est largement supérieur, en raison de l'effet de body flottant qui a un fort impact lorsque le MOS est bloqué. Le courant Ioff du BC régit par l'équation équivalente à celle du BULK est plus faible que le courant Ioff de la solution FB ou de la solution de MOS FB empilés pour une longueur de grille supérieure à 80nm. Dans cette dernière structure, le Vgs positif du transistor PMOS du bas réduit le courant sous le seuil et l'empilement assure une différence de potentiel Vds divisée par deux

sur chaque MOS. Le courant  $I_{off}$  de la solution de MOS FB empilés est assez constant car un effet de balancier existe entre le courant de canal sous le seuil et le courant de grille : quand le courant sous le seuil diminue avec l'augmentation de la longueur de grille, le courant de grille augmente, engendrant cette pseudo non-dépendance du courant de fuite total en fonction de la longueur de grille  $L$ . Ces résultats montrent qu'en SOI, c'est le MOS de type BC qu'il faut utiliser comme interrupteur de puissance avec une longueur de grille non minimale et selon le critère  $I_{on}/I_{off}$ .



**figure :  $I_{on}/I_{off}$  à 25°C et 125°C pour un interrupteur de puissance de type BC, FB et FB empilés pour une tension  $V_{DD}=1,2V$**

La figure présente une analyse en température de ces trois solutions. A 25°C la solution de MOS FB empilés a le ratio  $I_{on}/I_{off}$  le plus élevé, pour une longueur de grille minimale de 60nm. Par contre, pour des longueurs de grille  $L$  supérieures à 60nm ce ratio chute (le courant  $I_{on}$  est réduit plus fortement que le courant  $I_{off}$ ) et celui du BC est bien meilleur avec un optimum pour une longueur de grille compris entre 180 et 240nm. Pour ce point de fonctionnement, il est meilleur d'une décade par rapport aux FB empilés. Concernant l'interrupteur de puissance FB, son ratio  $I_{on}/I_{off}$ , est inférieur de deux décades par rapport à celui du BC en raison d'un courant  $I_{off}$  trop important.

A 125°C, l'écart entre la solution FB empilés et la solution BC se réduit, car le courant  $I_{off}$  croît plus fortement en BC avec la température. Le rapport  $I_{on}/I_{off}$  de la solution FB reste le moins bon. Le courant  $I_{off}$  étant plus élevé à 125°C, le ratio  $I_{on}/I_{off}$  de la solution BC a son optimum pour la longueur de grille  $L=240nm$ . Il chute pour une longueur supérieure à cause du courant de grille qui augmente. Le courant de fuite  $I_{off}$  est plus fortement réduit que le courant  $I_{on}$  en augmentant la longueur de grille. Cela se retrouve dans l'augmentation du rapport  $I_{on}/I_{off}$ . Globalement, avec une longueur de grille non minimale  $L=240nm$ , le BC a le rapport  $I_{on}/I_{off}$  le plus intéressant en fonction des deux températures étudiées.

Notons qu'il existe une étude où la longueur ou la largeur de grille de ces deux MOS empilés n'est pas la même [Das'03]. Ainsi, le transistor PMOS du bas peut avoir une longueur ou largeur différente de celui du haut car il a un  $V_{gs}$  positif en mode éteint et donc moins de courant de fuite. Cette étude a été réalisée en 65nm PD-SOI. Elle n'a pas montré de résultats convaincants permettant d'améliorer le ratio  $I_{on}/I_{off}$  de la solution FB empilés.

### 3.2.2.3. Optimisation utilisant une polarisation du body des interrupteurs de puissance

En considérant le MOS BC comme la meilleure solution d'interrupteur de puissance, une optimisation est proposée en réalisant une polarisation statique ou dynamique du body. La polarisation statique est contrôlée par une unité de gestion de la puissance via une alimentation extérieure VDDS, alors que la polarisation dynamique est contrôlée directement par le signal de contrôle de grille des interrupteurs de puissance.

Afin d'augmenter le ratio  $I_{on}/I_{off}$ , une polarisation en direct FBB est appliquée sur le potentiel du body afin de réduire la tension de seuil  $V_{th}$  et la résistance  $R_{on}$  du transistor et donc d'augmenter le courant  $I_{on}$ . A l'inverse, une polarisation de body inverse RBB est appliquée sur le body afin de réduire le courant de fuite  $I_{off}$  en augmentant la tension seuil  $V_{th}$  et donc  $R_{off}$ . De manière usuelle, en 65nm PD-SOI et BULK, cette polarisation statique contrôlée par une alimentation extérieure est limitée, pour un NMOS, à -300mV pour le RBB et à 300mV pour le FBB. Pour une valeur, inférieure à -300mV pour le RBB, le courant inverse de jonction devient supérieur au courant de canal sous le seuil ; alors que pour un FBB au-delà de 300mV, vers 600mV, la diode body/source devient fortement polarisée en direct. Le tableau montre le gain en  $I_{on}$  avec une polarisation FBB de 300mV à 25°C et la réduction de  $I_{off}$  avec une polarisation RBB de -300mV à 125°C (le courant de fuite  $I_{off}$  est plus élevé à 125°C). Le ratio  $I_{on}/I_{off}$  croisé ( $I_{on}$  à 25°C et  $I_{off}$  à 125°C) s'améliore de 88%.

Polarisation	$I_{on}$ avec FBB	$I_{off}$ avec RBB	$I_{on}/I_{off}$
Gain en courant	+9,1%	-42%	+88%

**tableau : Optimisation du ratio  $I_{on}/I_{off}$  de l'interrupteur de puissance BC,  $L=200nm$ , à l'aide de la polarisation de body**

L'inconvénient de la polarisation statique est la complexité de gestion du potentiel sur le body. Cela nécessite notamment une ou plusieurs Entrées/Sorties (I/O) supplémentaires, ou

des alimentations internes. La polarisation dynamique contrôlée par le signal de grille des interrupteurs de puissance permet de réduire cette complexité.

### 3.2.3. Optimisation de la surface d'implémentation

Comme précisé en début de chapitre, le but du réseau d'interrupteurs de puissance est de couper le courant de fuite en proposant une forte résistance équivalente  $R_{off}$ , tout en apportant une chute de tension «  $V_{drop}$  » faible, donc un faible  $R_{on}$ . Pour cela il suffit de réduire ou d'augmenter le nombre d'interrupteurs de puissance unitaires,  $N_{SW}$ .

Les stratégies d'optimisation proposées dans la littérature s'attachent prioritairement à diminuer la valeur de  $R_{off}$  [Das'03] pour une chute de tension donnée  $V_{drop}$  [Hwa'06], engendrant une faible dégradation du délai. Or la surface d'implémentation est une contrainte industrielle de plus en plus forte et le nombre d'interrupteurs de puissance à implanter dans le circuit, pour une chute de tension donnée, devient de plus en plus important, suite à l'augmentation du courant  $I_{on}$  lié à la réduction des technologies.

Si deux solutions apportent le même  $I_{on}/I_{off}$ , la meilleure sera celle qui nécessite le moins de surface d'implémentation pour le même  $R_{on}$ .

A titre d'illustration, pour l'interrupteur de puissance BC, l'intérêt est de connaître pour des longueurs de grille allant de 60nm à 240nm, laquelle est vraiment la plus intéressante. En considérant uniquement le ratio  $I_{on}/I_{off}$ , à 25°C les résultats sont très proches, mais à 125°C mieux vaut une longueur de grille élevée. Or ce raisonnement n'inclut pas la surface d'implémentation nécessaire. C'est pourquoi l'un des résultats de cette thèse est de proposer un nouveau facteur de sélection qui inclut, en plus des paramètres électriques  $R_{on}$  et  $R_{off}$ , la surface d'implémentation des interrupteurs de puissance pour une même chute de tension «  $V_{drop}$  ».

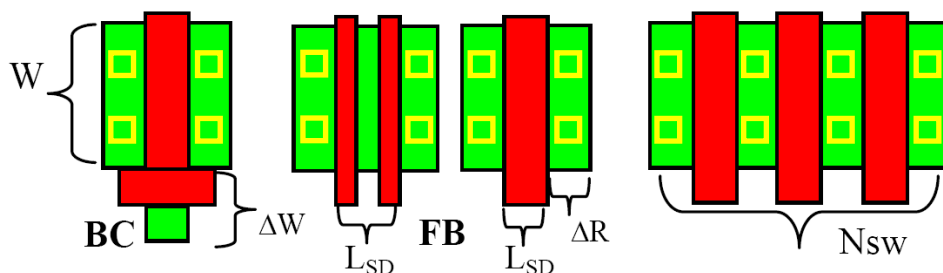


figure : Principe d'aboutement d'un réseau d'interrupteur de puissance

La figure illustre les cas étudiés : le motif élémentaire est constitué soit d'un transistor à body contacté, soit d'un ou deux transistors à body flottant empilés. Ce motif est ensuite NSW fois abuté dans une structure type multi-doigts. La surface nécessaire à l'implémentation du réseau d'interrupteurs de puissance (Area), dépend de plusieurs paramètres (équation ).

$$\text{Area} = \text{NSW} * (\text{LSD} + \Delta R) * (W + \Delta W)$$

**équation**

NSW est le nombre de transistors ou d'interrupteurs de puissance abutés ; LSD est la distance entre le drain (VDDI) et la source (VDD) de l'interrupteur de puissance unitaire ;  $\Delta R$  est la distance entre deux grilles (c'est la zone active qui permet l'insertion de contacts) ; W est la largeur et  $\Delta W$  le supplément de largeur lié à la prise body. LSD est la longueur de grille dans les cas BC et FB, mais vaut deux fois la longueur de grille plus la distance entre ces deux grilles sans contact, pour les solutions d'interrupteurs de puissance FB empilés. En 65nm, la distance  $\Delta R$  vaut 200nm alors que la distance entre deux grilles sans contact vaut 130nm. La surface est donc la même pour un interrupteur unitaire FB d'une longueur de grille de 250nm et un interrupteur de puissance unitaire composé de deux MOS FB empilés d'une longueur de grille de 60nm chacun et distants de 130nm.

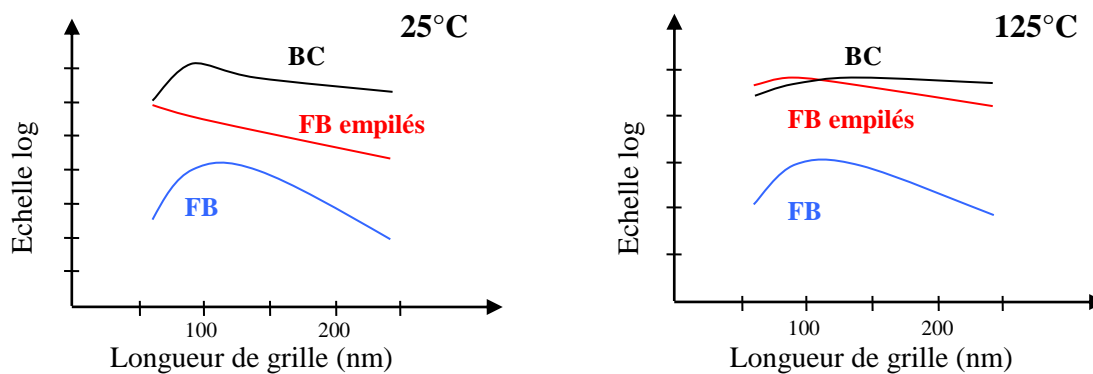
Il reste à savoir quel est le nombre d'interrupteurs de puissance NSW, à implémenter dans le réseau en fonction de la chute de tension  $V_{\text{drop}} = f(R_{\text{on}})$  maximale autorisée. Car, c'est ce paramètre qu'il faut respecter en priorité. En effet, au-delà d'une chute de tension spécifiée, le risque d'avoir une dégradation de délai non souhaitée apparaît.

Le nouveau facteur de mérite, proposé (équation ), prend en compte l'importance de la surface.  $I_{\text{on}}/\text{Area}$  est inversement proportionnelle à  $R_{\text{on}} * \text{Area}$ . Plus le rapport  $I_{\text{on}}/\text{Area}$  est élevé, plus le nombre d'interrupteurs de puissance unitaires (NSW) et donc la surface d'implémentation, est faible, pour une même chute de tension  $V_{\text{drop}}$  fixe. Notons qu'à partir d'un certain nombre de transistors implémentés, NSW,  $I_{\text{on}}/\text{Area}$  devient une quasi-constante.

$$\left( \frac{I_{\text{on}}}{I_{\text{off}}} \right)^X \times \left( \frac{I_{\text{on}}}{\text{Area}} \right)^Y$$

**équation**

Dans cette équation, X et Y sont respectivement des poids donnés aux paramètres  $I_{on}/I_{off}$  et  $I_{on}/Area$ .  $I_{on}/I_{off}$  est le ratio entre le courant ON et le courant de fuite OFF.  $I_{on}/Area$  est le courant ON par unité de surface. Libre au concepteur de choisir une valeur X supérieure à Y pour donner plus d'importance au ratio  $I_{on}/I_{off}$  tout en considérant  $I_{on}/Area$  ; ou de fixer  $X=Y=1$  afin de donner autant d'importance à  $I_{on}/I_{off}$  et à la surface d'implémentation. Si la surface n'est pas une contrainte alors Y peut être égale à 0. Ce critère permet néanmoins d'écarter la solution la plus couteuse en surface, dans le cas où les deux solutions les plus avantageuses auraient le même  $I_{on}/I_{off}$ .



**figure : Résultat du facteur de mérite des différentes solutions d'interrupteurs de puissance à 25°C et 125°C pour X=2 et Y=1**

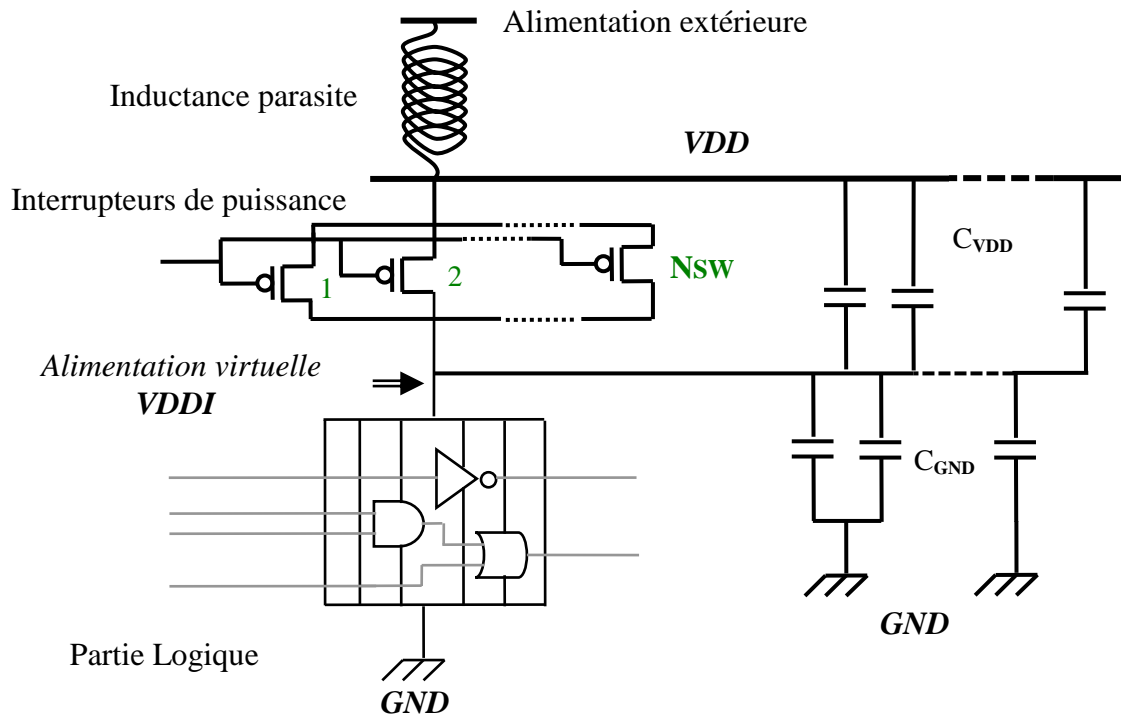
A l'aide de ce facteur de mérite, la solution BC pour une longueur de grille non minimale (L compris entre 90nm et 200nm) est la meilleure solution (figure ) car elle est la moins gourmande en surface. La valeur absolue de ce facteur de mérite n'est pas en soit une indication ; c'est par contre un critère de comparaison pertinent. L'échelle et l'unité des ordonnées des courbes de la figure ne sont pas importantes, elles varient en fonction des poids X et Y et peuvent être normalisées.

La longueur, L, des interrupteurs de puissance retenue selon ce critère et commune aux technologies 65nm BULK et PD-SOI est de 200nm. Elle remplace la longueur précédente L qui était fixée à 240nm.

#### 3.2.4. Capacités de découplage

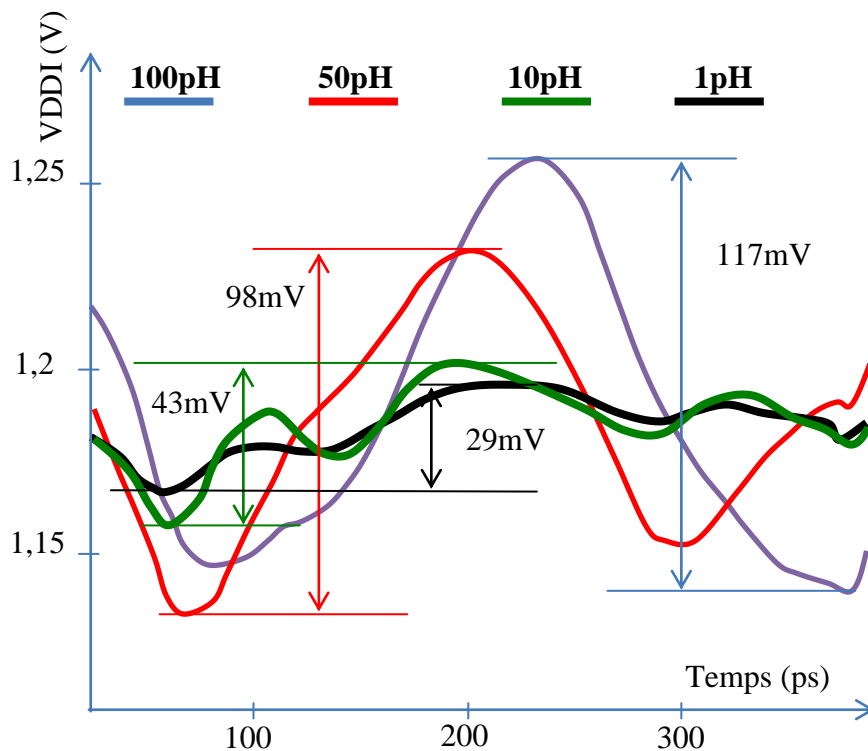
Pour filtrer les bruits de commutation, il est nécessaire d'introduire des capacités de découplage sur l'alimentation virtuelle (figure ). Ceci est encore plus important en technologie SOI, suite à l'absence du découplage naturel apporté par les capacités de jonction

caisson/substrat. Dans cette optique, tout espace libéré par les interrupteurs de puissance, peut être utilisé pour insérer des capacités de découplage, d'où l'importance du facteur de mérite proposé.



**figure : Technique MTCMOS avec capacités de découplage**

Les capacités de découplage peuvent être connectées à l'alimentation VDD ( $C_{VDD}$ ) ou à la masse GND ( $C_{GND}$ ). Ces capacités de découplage ont un rôle de lissage. En effet les appels de courant de la partie logique, et le fait que le réseau d'interrupteurs de puissance soit résistif et qu'il existe une inductance parasite entre l'alimentation extérieure et l'alimentation du circuit VDD, induisent une modulation de la tension de l'alimentation virtuelle VDDI. Les capacités de découplage permettent d'atténuer la chute de tension aux bornes de la résistance équivalente du réseau d'interrupteurs de puissance lors de ces appels de courant.



**figure : Fluctuation de la tension d'alimentation virtuelle VDDI d'un même circuit logique ayant les mêmes appels de courant en fonction de l'inductance parasite**

La figure montre que pour le même réseau d'interrupteurs de puissance ( $R_{on}$ ) et pour le même courant d'appel ( $I_{on}$ ), plus l'inductance parasite est grande, plus la fluctuation en tension de l'alimentation virtuelle VDDI est élevée. Dans ce cas aucune capacité de découplage n'est associée au circuit. L'intérêt d'implémenter ces capacités de découplage réside dans le fait de réduire cette fluctuation et ses conséquences. En effet ces fluctuations provoquent une dégradation des performances sur la partie logique et réduisent la fiabilité du circuit.

Considérons un réseau d'interrupteurs de puissance implémenté dans un circuit comprenant un nombre de transistor NSW, il requiert une surface PS. Si aucune capacité de découplage n'est associée au réseau d'interrupteurs alors la surface d'implémentation totale de la technique MTCMOS est  $T_A = PS$ . En conservant cette surface disponible  $T_A$ , et en remplaçant un certain nombre d'interrupteurs de puissance par des capacités de découplage, une analyse de la dégradation du délai et du courant de fuite  $I_{off}$  est proposée sur un bloc logique composé de plus d'un millier de transistors, et en fonction de plusieurs valeurs d'inductance parasite. Ce bloc logique a une signature en courant qui est à l'image de celles observées sur la plupart des blocs logiques synthétisés. L'intérêt de cette étude est de connaître la dégradation du délai par rapport à la réduction du courant de fuite en fonction des



capacités de découplage et des inductances parasites. Le nombre d'interrupteurs, et donc la surface PS, passe de 100% à 75% et 50% de la surface totale TA. Cela permet d'introduire respectivement 25% et 50% de capacités de découplage. Cette réduction du nombre d'interrupteurs de puissance va permettre de diminuer également le courant de fuite  $I_{off}$ , mais aussi la conductance du réseau,  $1/R_{on}$ , de 25% et 50%.

PS/TA	Inductance	50pH	100pH
100%		1	1,28
75%		1,01	1,19
50%		1,03	1,07

**tableau : Produit Délai de dégradation  $\times$  Ioff en fonction du ratio PS/TA pour différentes valeurs d'inductance parasites**

Le tableau montre le produit Dégradation du délai ( $D_d$ )  $\times$  Courant de fuite ( $I_{off}$ ) en fonction du rapport entre les surfaces d'interrupteurs de puissance et totale, et en fonction de l'inductance parasite. Le transistor utilisé est du type BC avec une longueur de grille  $L=100\text{nm}$  et une polarisation nulle,  $V_{bs}=0$ .

A température ambiante, et pour un circuit logique générique composé de plusieurs portes logiques, ayant un appel en courant représentatif, il est visible que plus l'inductance parasite est grande, plus il est intéressant d'introduire des capacités de découplage. En effet, pour une faible inductance parasite l'impact sur le délai reste faible. En revanche, pour de fortes inductances parasites la valeur de la fluctuation d'alimentation est plus grande, induisant une dégradation du délai plus élevée. Introduire des capacités de découplage assure une réduction de la dégradation de délai lorsque l'inductance parasite est grande car elles permettent de lisser la tension d'alimentation et d'être moins sensible aux appels en courant,  $I_{on}$ . Elles n'améliorent pas le rapport  $I_{on}/I_{off}$  des interrupteurs auxquelles elles sont associées mais réduisent la chute de tension  $V_{drop}$  et donc la dégradation de délai. Cela démontre que la dégradation de délai n'est pas linéairement dépendante du courant  $I_{on}$  qu'est capable de fournir le réseau d'interrupteurs de puissance mais plutôt de  $V_{drop}$  en fonction du temps. Les éléments parasites du réseau de distribution d'alimentation forment donc un élément supplémentaire à considérer en technologie PD-SOI.

Une optimisation d'une solution en termes de  $I_{on}/I_{off}$  peut se faire soit sur la réduction de  $I_{off}$  soit par l'amélioration de  $I_{on}$ . A surface égale l'amélioration de  $I_{on}$  permet de réduire le nombre d'interrupteurs de puissance ( $N_{SW}$ ), et donc d'implémenter plus de capacités de découplage ; d'où l'intérêt de proposer un facteur de sélection qui prend en compte la surface d'implémentation.

### 3.3. Proposition d'optimisation du Ron de l'interrupteur de puissance BC

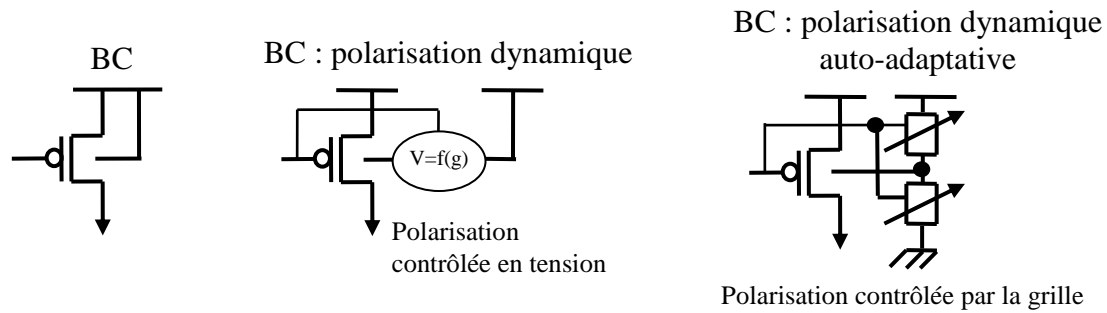
#### 3.3.1. Présentation du montage Auto-DTMOS proposé

Il est donc préférable de proposer une solution d'interrupteurs de puissance ayant un  $I_{on}$  25% plus élevé plutôt qu'un  $I_{off}$  20% plus faible car cela permet de réduire la surface des interrupteurs de puissance de 20%. Le rapport est en effet le même :  $(I_{on} \times 1,25) / I_{off} = I_{on} / (I_{off} \times 0,8)$ , mais comme  $I_{on}/Area$  est 25% supérieure, cette solution est meilleure selon le facteur de sélection proposé (équation ). En optimisant la solution BC et en proposant une polarisation directe du body, il est possible de réduire la valeur de la résistance Ron.

La polarisation de body est une solution gourmande en surface quand elle nécessite un bloc I/O d'alimentation (Entrée/Sortie) dédié et supplémentaire. La solution proposée devra conserver la même complexité pour une surface égale ou réduite ; ainsi proposer une polarisation statique à l'aide d'une alimentation extérieure doit être évitée. L'introduction d'un bloc I/O d'alimentation de body induit, en plus d'une surface supplémentaire, une gestion plus complexe. La solution proposée part du principe que l'optimisation doit garantir au concepteur un possible échange avec la solution initiale : BC avec polarisation nulle ( $V_{bs}=0$ ).

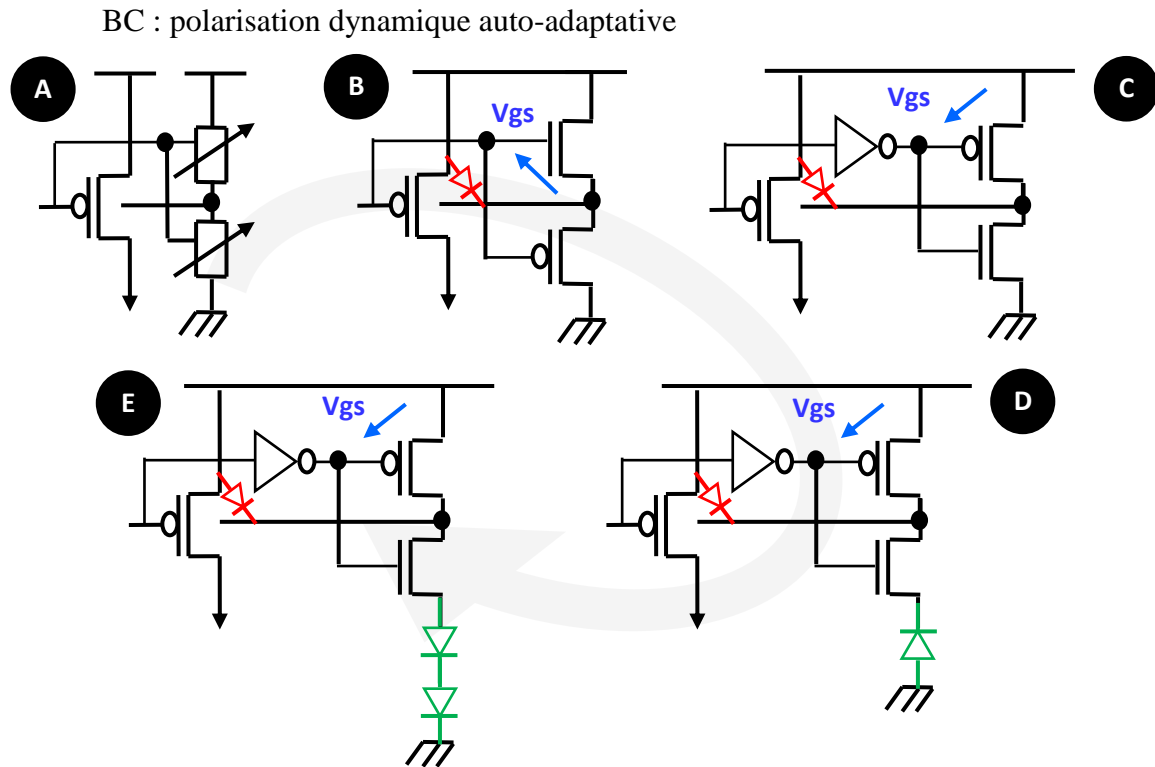
Pour optimiser le ratio  $I_{on}/I_{off}$  et donc la valeur de Ron, une polarisation dynamique et auto-adaptative est proposée. Cette solution ne nécessite pas de source d'alimentation de body supplémentaire, et elle permet de fonctionner avec le même contrôleur. En effet, une modification du contrôleur serait inévitable si celui-ci devait gérer, en plus, l'alimentation du body.

Comme cette proposition de polarisation dynamique et auto-adaptative de body amène une amélioration globale du Ron, elle nécessitera moins d'interrupteurs de puissance ; un certain nombre d'entre eux pourront donc être remplacés par des capacités de découplage, à surface constante.



**figure : Présentation du principe d'interrupteur de puissance optimisé à faible  $R_{on}$**

La figure présente le principe général de l'optimisation proposée, qui est basée sur un pseudo-pont diviseur dont la valeur des résistances est contrôlée par la grille de l'interrupteur de puissance. A la différence d'une solution BC avec polarisation dynamique où la tension de grille du transistor génère également une tension pour le body, et où le bloc connecté au body sert surtout de limiteur de courant, cette nouvelle solution est basée sur un potentiel auto-adaptatif qui ne requiert pas d'alimentation supplémentaire. Les montages utilisant des limiteurs et présentés dans le chapitre précédent ne permettent pas une polarisation de body suffisamment élevée. Il faut, pour polariser en direct le body au-delà de plusieurs centaines de millivolts et donc avoir une action efficace sur la tension de seuil, un montage fondé sur un pont diviseur. Alors que les limiteurs amènent une polarisation ayant un ordre de grandeur de la dizaine de millivolts, impactant peu la résistance  $R_{on}$ , un montage de polarisation en direct permet de réduire de manière significative la valeur de la résistance  $R_{on}$ .



**figure : Optimisation de l'interrupteur de puissance à faible  $R_{on}$  en 65nm PD-SOI**

La conception du montage qui permet de polariser en direct le body du réseau d'interrupteurs de puissance, afin de réduire sa résistance équivalente  $R_{on}$ , se fait par étape. A partir de la solution initiale proposée figure (A), la première étape pour obtenir une polarisation auto-adaptative de body en fonction du mode de fonctionnement des interrupteurs de puissance consiste à remplacer les résistances variables, commandées par le potentiel de grille, par le composant le plus simple ayant ce rôle : le MOS (B). Malheureusement, le type de connexion fait en (B) entraîne une chute de tension  $V_{gs}$  entre l'alimentation et la prise body représentée par une flèche bleue. Cette chute de tension ramène sur le body un potentiel inférieur à VDD alors que le réseau d'interrupteur est bloqué. Cela correspond à un léger FBB sur le transistor, valant  $V_{gs}$ , non-optimal pour réduire le courant de fuite  $I_{off}$ . Il suffit pour supprimer cette chute de tension de revenir à une configuration standard avec les MOS de type P connectés à l'alimentation VDD et les MOS de type N à la masse GND ; le signal de grille est donc inversé (C). Cependant, le body passe de VDD en mode OFF à GND en mode ON, ce qui met la diode Source/Body fortement polarisée en direct avec un potentiel à ses bornes valant VDD. Pour remédier à ce problème, un limiteur de courant, qui n'est ni plus ni moins qu'une diode en mode inverse est alors placé entre le NMOS et la masse GND (D). L'insertion de la diode amène un potentiel de body FBB en mode ON trop faible ; en effet la

diode inverse étant beaucoup plus résistive que la diode directe Source/Body, le pont diviseur amène un potentiel auto-adaptatif sur le body proche de VDD. Dans cette configuration le FBB conduit à une réduction de  $R_{on}$  à peine inférieure à 5% ; cette solution est donc écartée. Pour régler ce problème, nous avons proposé d'empiler plusieurs diodes en direct à la place de la diode en inverse (E) entre la masse et la source du NMOS, contrôlé par l'inverseur. Ceci crée ainsi le pseudo pont diviseur qui tire vers le bas, de manière auto-adaptative, le potentiel de « body », ce qui permet de réduire  $R_{on}$  de plus de 20%. Le comportement est auto-adaptatif, car, plus le courant du body est élevé, plus la tension aux bornes des 2 diodes en direct est élevée et plus la différence de potentiel aux bornes de la diode Source/Body s'abaisse, réduisant le courant de body. C'est un système rebouclé et stable qui assure une polarisation de body en direct. Il existe donc un effet de compensation et d'équilibre entre le courant et le potentiel de body. Cette solution appelée Auto-DTMOS, associée au réseau d'interrupteurs de puissance de type BC une structure de polarisation de body en direct, auto-adaptative et dynamique.

### 3.3.2. Implémentation du montage Auto-DTMOS en PD-SOI 65nm

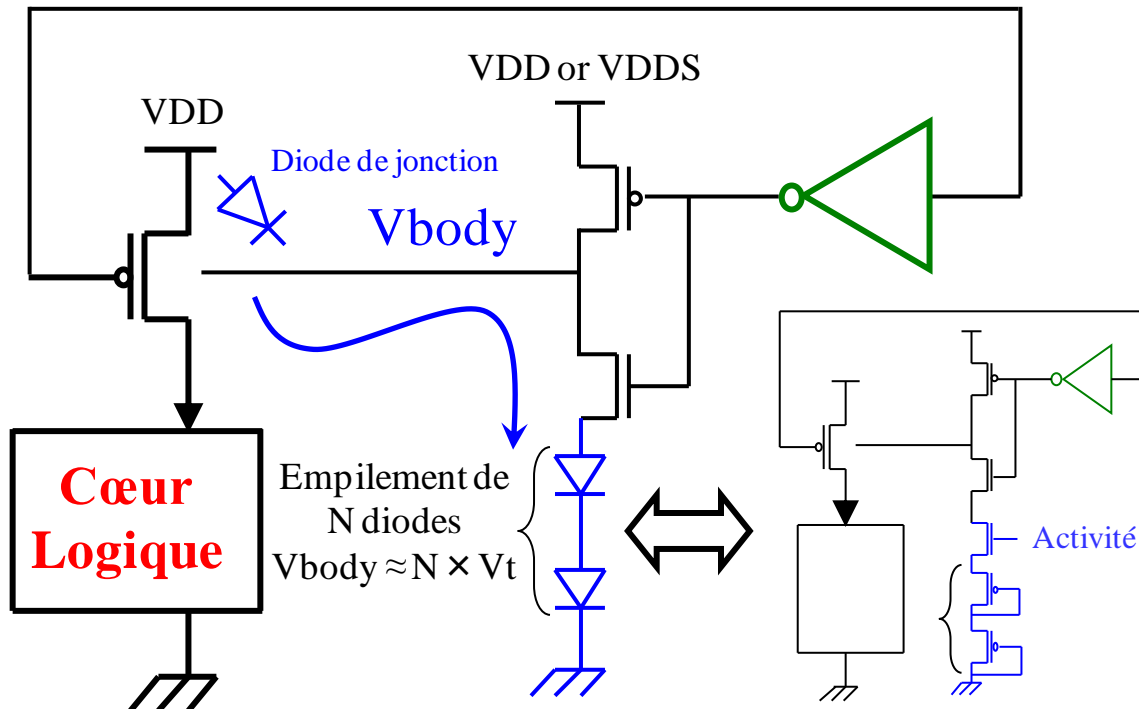


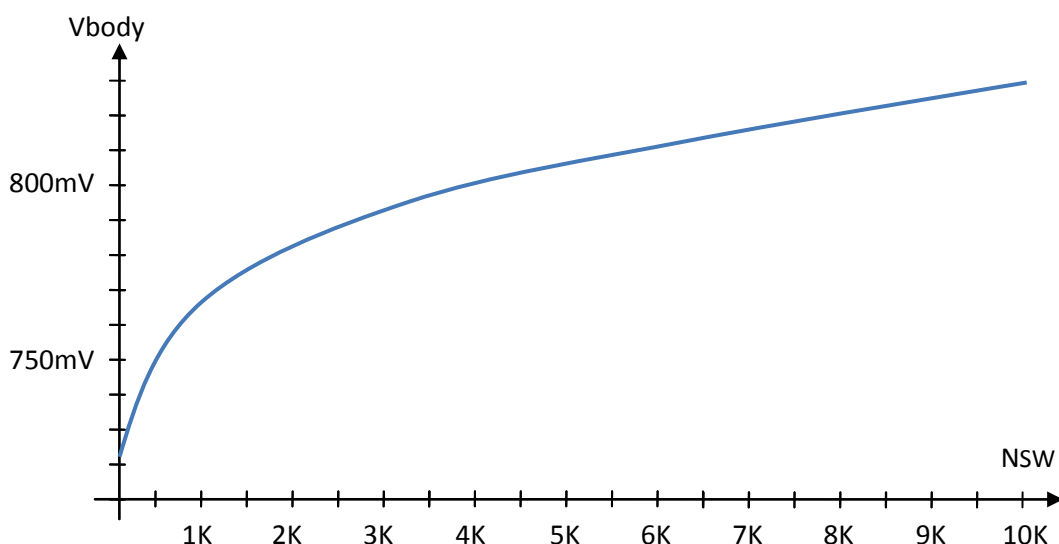
figure : Solution Auto-DTMOS à faible  $R_{on}$  en 65nm PD-SOI

La figure décrit la solution Auto-DTMOS brevetée, incluant le bloc de polarisation auto-adaptatif. Une ou plusieurs (N) diodes peuvent être empilées, créant un potentiel auto-

adaptatif  $V_{body}$  proche de  $N \times V_t$ , où  $V_t$  est la tension de seuil ou de coude de la diode. L'optimum est un empilement de deux diodes car il permet, en mode ON, un FBB plus intéressant qu'un empilement de trois diodes. Avec une seule diode utilisée, la polarisation directe est plus élevée mais le courant de body devient trop important. Ces diodes sont des MOS montés en diode avec la grille connectée au drain.

Une option supplémentaire consiste à insérer un MOS de contrôle d'activité. Dans le cas où le circuit est en mode ON, mais en période d'inactivité ; il est possible de gagner en consommation, en annulant le FBB et le courant allant du « body » à GND, même si celui-ci est très faible par rapport à  $I_{ddq}$ . Lorsque le niveau du signal « Activité » est bas, le potentiel de « body » des interrupteurs de puissance est flottant et proche de celui d'un transistor à body flottant. Le FBB, dans ce cas, est faible, mais comme le courant d'alimentation du circuit se limite au courant statique  $I_{ddq}$  et qu'il est plus faible de plusieurs décades que le courant en fonctionnement actif  $I_{on}$ , avoir un faible FBB et donc une résistance  $R_{on}$  plus élevée de quelques pourcents n'induit quasiment pas de chute de tension supplémentaire  $V_{drop}$ , et est sans conséquences.

Le bloc auto-adaptatif connecté au « body » de l'interrupteur de puissance peut être conçu et dimensionné en fonction du nombre total d'interrupteurs, NSW, qu'il aura à piloter. Dans une bibliothèque mieux vaut le concevoir comme un bloc indépendant et le multiplier dans le circuit en fonction du nombre NSW.



**figure : Tension de body ( $V_{body}$ ) d'un réseau composé de NSW interrupteurs de puissance polarisé par 1 seul bloc de polarisation de body auto-adaptatif**

La polarisation de body est la différence de potentiel entre le body et la source. Considérant la figure , et au vu du dimensionnement proposé pour le bloc de polarisation auto-adaptatif, le meilleur compromis semble être l'insertion d'un bloc de polarisation tous les 1000 / 2000 interrupteurs de puissance unitaires qui vont composer le réseau d'interrupteurs de puissance. Au-delà, la polarisation directe de body FBB des interrupteurs de puissance de type PMOS diminue puisque le potentiel de body,  $V_{body}$ , augmente et que la source reste à  $V_{DD}=1.2V$ , ce qui ne permet plus une réduction de  $R_{on}$  de 20%. En dessous de ce nombre, la perte relative de surface, devient trop importante. De plus si le nombre de blocs utilisés devient trop important, le courant de diode Source/Body du réseau d'interrupteurs de puissance devient très élevé, ce qui va réduire le gain en consommation statique. Apporter une polarisation FBB au moins supérieure à 400mV permet d'obtenir un gain en  $R_{on}$  d'au moins 20%, et semble un bon compromis.

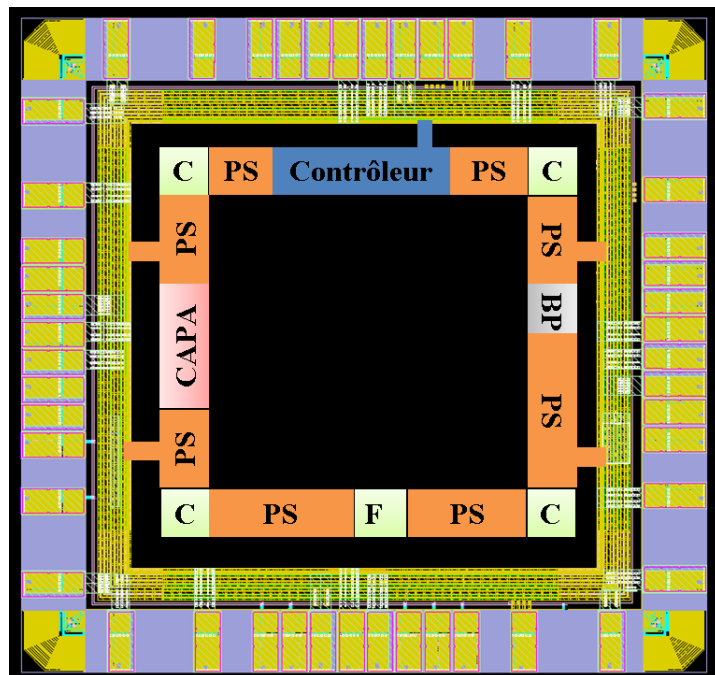
La solution retenue est donc l'intégration d'une structure de polarisation auto-adaptative environ tous les 1000 interrupteurs unitaires. L'utilisation de cette structure de polarisation du body permet de réduire de 20% le nombre d'interrupteurs unitaires implémentés, NSW. Or, la surface d'un bloc de polarisation auto-adaptatif est équivalente à la surface de cinq interrupteurs de puissance unitaires abutés, représentant donc 0,5% de la surface du réseau d'interrupteurs de puissance pilotable. La mise en œuvre de la structure proposée permet donc, en la comparant à une structure BC de référence :

- De réduire la surface de silicium occupée, à  $R_{on}$  constant, de 19,5%. Dans ce cas, comme le courant de fuite par unité de réseau reste le même, le courant de fuite total est également 20% plus faible.
- De réduire de 20% la résistance  $R_{on}$  typique à 25°C et  $V_{DD}=1.2V$ , pour le même nombre d'interrupteurs de puissance, NSW.

#### 3.3.3. Développement de la bibliothèque associée au montage Auto-DTMOS et Réalisation du circuit de test

Une bibliothèque complète MTCMOS, basée sur la structure Auto-DTMOS, ainsi qu'un circuit de validation et de comparaison ont été développés (figure ). Ce circuit de test permet de faire les principales mesures pour le réseau d'interrupteurs de puissance, qui sont les résistances équivalentes ( $R_{on}$  et  $R_{off}$ ), le temps de réveil, la consommation dynamique et statique, ainsi que la validation du contrôle et des différents modes.

« PS » est le réseau d'interrupteurs de puissance ou « power switch » ; « C » sont les coins, qui permettent de changer l'orientation du réseau et ainsi de créer un anneau ; « F » sont les cellules de remplissage qui permettent d'assurer la continuité électrique tout comme les coins ; CAPA est la capacité de découplage et « BP » est la structure de polarisation auto-adaptative.

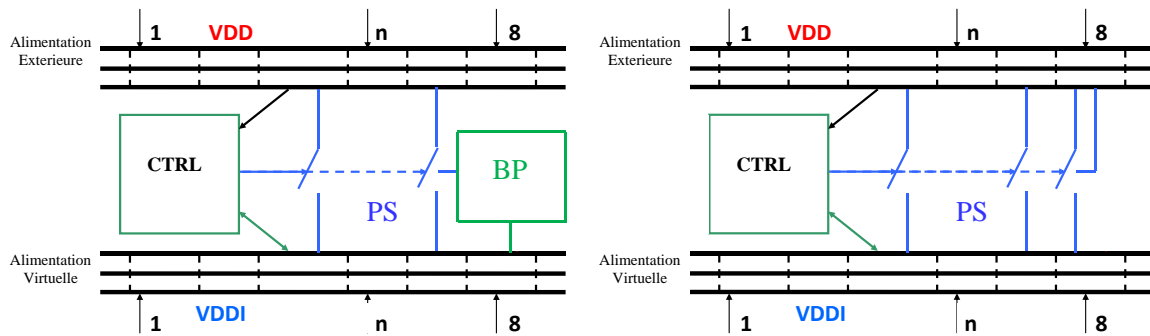


**figure : Circuit de test de la librairie MTCMOS en anneau**

Dans ce circuit de test, le drain, la source, la grille et le body du réseau d'interrupteurs de puissance sont accessibles permettant d'effectuer toutes les mesures possibles de la solution nouvelle proposée.

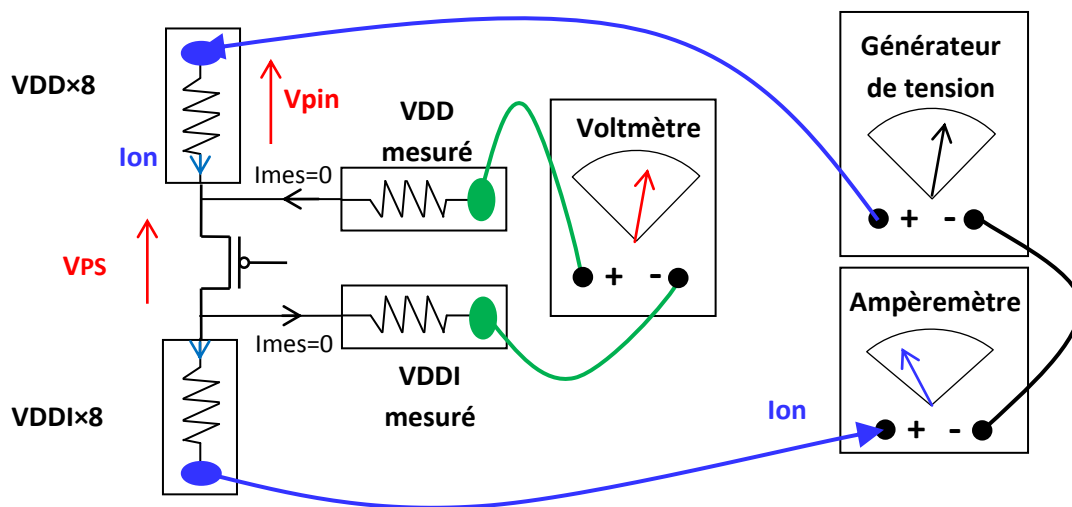
Deux autres variantes de ce circuit de comparaison ont été conçues. Le second circuit a été conçu en PD-SOI et implémenté avec la solution initiale : des interrupteurs de puissance de type BC à polarisation nulle (le « body » étant relié à VDD). Le troisième circuit a été conçu en BULK et porté en PD-SOI ; ainsi tous les interrupteurs de puissance sont de type FB. Tous ces circuits permettent de valider toutes les solutions et de réaliser des comparaisons. Dans les trois circuits, le nombre d'interrupteurs de puissance en parallèle est le même, NSW=3500. Le circuit de test avec la solution Auto-DTMOS intègre en plus 4 blocs de polarisation de body auto-adaptative, « BP », un tous les 875 interrupteurs de puissance unitaires.





**figure : Schéma électrique de l'alimentation des circuits de test de la bibliothèque associée au montage Auto-DTMOS et de la bibliothèque initiale**

La figure montre le schéma électrique équivalent permettant de mesurer  $R_{on}$  et  $R_{off}$  des circuits de test de la bibliothèque associée au montage Auto-DTMOS et à la bibliothèque initiale (interrupteurs de puissance de type BC avec polarisation de body nulle). Les modes ON, OFF, de mise en veille et de réveil sont gérés par les contrôleurs respectifs de chaque circuit, CTRL. L'accès direct à l'alimentation extérieure, VDD, et l'alimentation virtuelle, VDDI, permettent ces mesures. Dans ce circuit de test il existe huit plots pour chaque alimentation VDD et VDDI.



**figure : Schéma électrique de la mesure de  $R_{on}$  du réseau d'interrupteurs de puissance**

La figure montre le principe de mesure des paramètres  $R_{on}$  et  $R_{off}$ . Le courant traversant le réseau d'interrupteurs de puissance traverse également les 8 plots d'alimentation dédiés. Or, ces plots ont une résistance équivalente, du même ordre de grandeur que la résistance équivalente du réseau ( $R_{on}$ ) lorsque celui est activé. Ils vont donc entraîner une chute de tension  $V_{pin}$ , en raison du courant  $I_{on}$  qui les traverse, et qui fausse la mesure de la

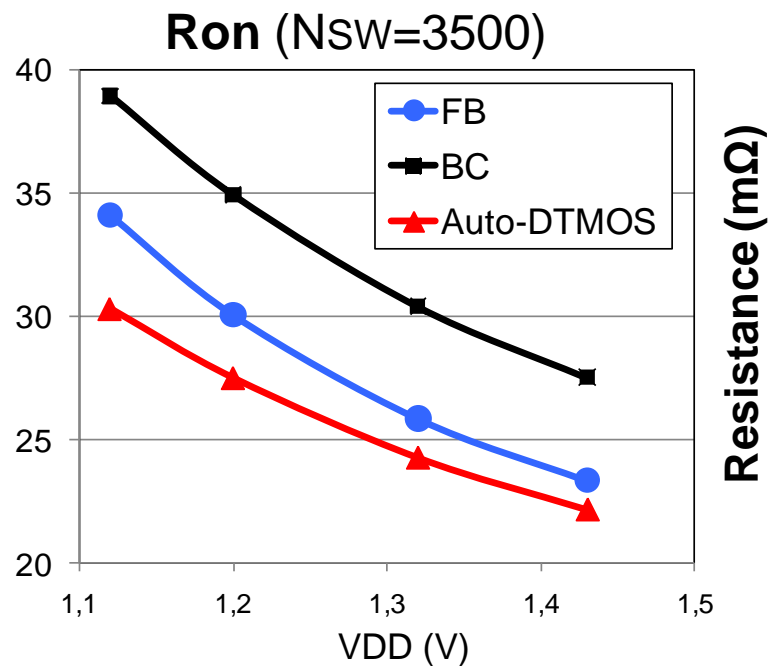
tension aux bornes des interrupteurs de puissance, VPS. Entre VDD et VDDI la tension est égale à  $VPS + 2 \times V_{pin}$ . C'est pour cela qu'une mesure précise à 4 points est faite, qui permet le calcul de  $R_{on}$  par la relation  $VPS/I_{on}$ .  $I_{on}$  est mesuré par l'ampèremètre connecté en série avec le générateur de tension délivrant la différence de potentiels entre VDD et VDDI. Pour Roff, la résistance du réseau d'interrupteurs de puissance étant largement supérieure aux résistances des plots d'alimentation, il n'est pas nécessaire de se mettre dans de telles conditions de mesures puisque  $VPS=VDD$ .

### 3.3.4. Mesures et comparaisons des interrupteurs de puissance FB, BC et Auto-DTMOS

La première mesure à faire est la validation fonctionnelle des 3 solutions. Pour cela le circuit doit pouvoir mettre en marche/arrêt le réseau d'interrupteurs de puissance et vérifier que le temps de réveil respecte un cahier des charges bien précis. La comparaison montre que les temps de réveil sont sensiblement les mêmes et qu'ils respectent bien ceux simulés lors de la conception. Pour mesurer ce temps de réveil, comme le circuit ne contient que la bibliothèque MTCMOS, il faut connecter une capacité de même ordre de grandeur que la capacité équivalente d'un circuit logique que l'alimentation virtuelle pourrait alimenter. La capacité équivalente d'un circuit varie quasi linéairement avec le nombre de portes que celui-ci contient. Le circuit de test comprend 3500 interrupteurs de puissance unitaires. En partant du courant qui peut être fourni par ces interrupteurs de puissance, nous avons estimé la taille du circuit qu'il serait possible d'alimenter et sa capacité totale équivalente. Deux mesures de réveil ont été réalisées avec deux capacités différentes : une maximale de 47nF et une moyenne valant 1nF. Pour des conditions de réveil optimales et pour la capacité moyenne de 1nF, le temps de réveil peut être réduit à 10µs. A l'opposé pour un circuit équivalent à une capacité de 47nF, et pour des conditions défavorable, le temps de réveil peut être plus long : environ de 500µs.

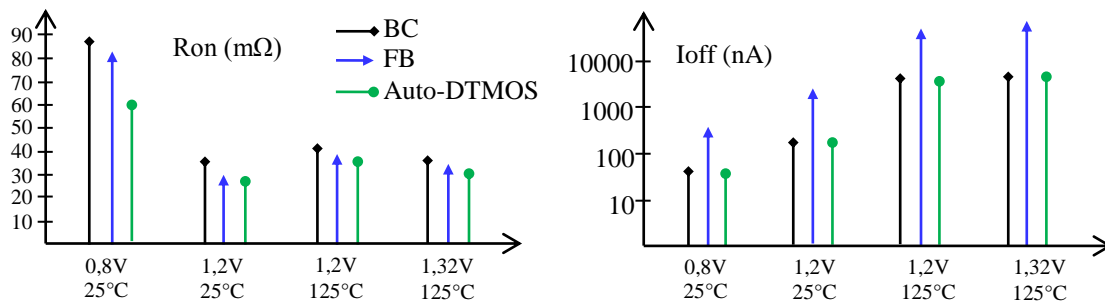
Le contrôleur agit sur le temps de réveil, en contrôlant la grille des interrupteurs de puissance, SWVDD, mais aussi en limitant le courant d'appel,  $I_{max}$ , lors de la phase de réveil [Sch'09]. Le contrôleur est dans un mode lent dans des conditions défavorables de temps de réveil et dans un mode rapide pour des conditions optimales de temps de réveil. La différence entre ces deux modes de réveil se fait sur la limitation du courant d'appel délivré à la partie logique et le potentiel de la grille des interrupteurs de puissance. Le contrôleur, dans le mode lent, limite le courant d'appel à une dizaine de microAmpères et amène lentement le potentiel

de grille des interrupteurs de puissance, SWVDD, de VDD (bloqué) à GND (passant) ; alors que dans le mode rapide, il peut libérer plus d'une centaine de microAmpères et rapidement faire commuter les interrupteurs de puissance. Le mode de réveil lent a l'avantage d'éviter un fort appel en courant.



**figure : Mesure sur silicium de la résistance équivalente Ron des 3 circuits de test des interrupteurs de puissance FB, BC et Auto-DTMOS**

La figure montre la mesure de Ron des interrupteurs de puissance de type FB, BC et Auto-DTMOS, pour plusieurs tensions d'alimentation VDD. La valeur de la résistance Ron du réseau constitué de la solution Auto-DTMOS est 20% plus faible que la valeur de la résistance du réseau de la solution initiale BC pour le même nombre NSW=3500 et pour une tension comprise entre 1,1V et 1,3V.



**figure : Mesure sur silicium de Ron et de Ioff du réseau d'interrupteurs de puissance FB, BC et Auto-DTMOS pour plusieurs points significatifs**

La résistance du réseau FB est d'environ 10% plus faible que celle du BC, tout en restant supérieure à celle de la solution Auto-DTMOS. Cependant, le courant Ioff du réseau FB est supérieur de plus d'une décade par rapport au deux autres cas et ceci pour plusieurs cas significatifs de températures et tensions d'alimentation (figure ). La résistance équivalente du réseau d'interrupteurs de puissance passant, est la plus élevée pour la valeur VDD minimale. Dans ce cas la résistance de la solution Auto-DTMOS est même 30% plus faible que celle de la solution BC, pour le même nombre d'interrupteurs de puissance unitaires, et à 25°C. C'est ce cas de figure qui introduit le plus de dégradation de délai sur le circuit et qui amène le délai le plus lent sur la partie logique.

Le point de fonctionnement VDD=0,8V, température=125°C et une qualité de silicium lent est utilisé pour déterminer, pour un circuit logique, le nombre d'interrupteurs de puissance unitaires à utiliser, NSW, et la chute de tension « Vdrop » maximale autorisée. Pour ce point de fonctionnement, le gain en Ron est inférieur à 30% ; il permet de réduire de 20% le nombre d'interrupteurs sans problème.

Pour une valeur VDD supérieure à la tension nominale (VDD=1,35V), à 125°C, et pour le même nombre d'interrupteurs de puissance (NSW), la résistance équivalente Ron de la solution Auto-DTMOS est 15% plus faible que la solution BC. Pour ces valeurs de tension supérieures à 1,2V, la vitesse du circuit est la plus grande et la résistance du réseau d'interrupteurs est la plus faible.

Dans les cas typiques et pires cas, nous avons la possibilité de réduire le nombre d'interrupteurs de puissance de 20% grâce à la solution Auto-DTMOS. Avec cette solution, on a, par rapport à la solution BC, une résistance Ron 6,25% plus élevée à VDD=1,35V et à 125°C. Mais pour ces valeurs de tensions, la vitesse du circuit est la plus élevée, et une dégradation de la résistance et de la chute de tension est donc sans importance sur la tenue des spécifications. Par contre, à plus basse tension d'alimentation (VDD=0,8V) et pour un

nombre d'interrupteurs de puissance unitaires toujours 20% plus faible, la valeur de  $R_{on}$  est 12,5% plus faible, entraînant une chute de tension et donc une dégradation du délai moindre. La solution Auto-DTMOS offre donc un fonctionnement plus avantageux dans le cas le plus défavorable en termes de performances, à faible VDD. Ceci permettrait d'être encore plus agressif sur le nombre d'interrupteurs de puissance à implémenter si la conception et la validation du circuit se faisaient à ce point de fonctionnement.

Le courant de fuite,  $I_{off}$ , est identique dans les cas du BC et de l'Auto-DTMOS, pour le même nombre NSW, car tous deux ont une polarisation de body nulle,  $V_{bs}=0$ . A même  $R_{on}$  visée, avec 20% d'interrupteurs de puissance en moins, la solution Auto-DTMOS a un courant de fuite ainsi qu'une surface d'implémentation 20% plus faible. Les mesures sur silicium ont montré par contre, que le courant de fuite  $I_{off}$  de la solution FB est trop élevé, en raison de l'effet de body flottant qui abaisse sa tension de seuil  $V_{th}$ . La solution FB doit donc être clairement écartée comme solution d'interrupteur de puissance au profit de la solution BC ou Auto-DTMOS.

#### 3.4. Conclusion sur l'Auto-DTMOS

Dans ce chapitre, une méthode de dimensionnement, s'appuyant sur un facteur de mérite, a été proposée. Il apparaît qu'en 65nm LP PD-SOI, l'interrupteur de puissance optimal est de type « body contacté » (BC). La solution Auto-DTMOS, basée sur une polarisation auto-adaptative du body d'interrupteurs de puissance BC a été présentée. Cette solution validée sur silicium permet un gain en surface et courant de fuite de -20%, permettant ainsi d'inclure plus de capacités de découplage. Cette solution a été retenue pour la suite de ce travail. Dans le prochain chapitre son intégration dans un circuit démonstrateur avec une partie logique complexe de type LDPC est détaillée.

- [Das'03] Das, K.K. et al. "New Optimal Design Strategies and Analysis of Ultra-Low Leakage Circuits for Nano-Scale SOI Technology" ISLPED, 2003
- [Fla'09] Flatresse, P.; Le coz, J. ; Raynaud, C. ; Thomas, O. "Digital SOI Design for Low Power Applications" MIGAS, 2009
- [Hwa'06] Hwang, C.; Kang, C. & Pedram, M "Gate sizing and replication to minimize the effects of virtual ground parasitic resistances in MTCMOS designs" Quality Electronic Design, 2006
- [Kea'07] Keating, M. et al. "Low Power Methodology Manual For System-on-Chip Design" Springer, 2007
- [Lho'05] L'Hostis, N.; Valentian, A. & Amara, A. "A 130nm partially depleted SOI technology menu for low-power applications" *IEEE NEWCAS Conference*, p.175-178, 2005
- [Qi'06] Qi, X. et al. "Efficient subthreshold leakage current optimization - Leakage current optimization and layout migration for 90- and 65- nm ASIC libraries" *IEEE CIRCUITS & DEVICES MAGAZINE*, p. 39-47, sept/oct 2006
- [Sch'09] Schoellkopf, J.-P.; Magarshack, P. "Low-Power Design Solutions for Wireless Multimedia SOCs" *Design & Test of Computers, IEEE*, vol. 26, p. 20-29, 2009
- [Tho'09] Thomas, O.; Belleville, M.; Liot, V. & Flatresse, P. "Method and Device for Adapting the Voltage of a MOS Transistor Bulk" Patent N°.: US 7,622,983 B2 , Nov. 24, 2009



## 4. Application à un circuit LDPC

*Dans ce chapitre, la réalisation d'un circuit démonstrateur incluant un bloc logique de haute complexité, appelé LDPC, pour « Low Density Parity Check », est présentée. Ce circuit, incluant la technique MTCMOS est dérivé en plusieurs versions : (1) en technologie BULK avec des interrupteurs de puissance classiques ;(2) en technologie PD-SOI avec des interrupteurs BC sans polarisation, mais aussi (3) avec la solution proposée Auto-DTMOS. Le but est de comparer ces différentes implémentations mais également de comparer les technologies BULK et PD-SOI – de type LP (Low-Power) et de nœud technologie 65nm –. La stratégie d'implémentation des interrupteurs ainsi que leur nombre sont précisés pour ce cas applicatif.*



#### 4.1. Description du circuit LDPC

Pour évaluer l'impact des interrupteurs de puissance sur le fonctionnement du circuit alimenté, tant en termes de fréquence de fonctionnement que de réduction des courants de fuite, il convient de choisir un circuit applicatif relativement complexe, représentatif de ce qui peut se retrouver dans des applications portables actuelles. Ainsi, le choix s'est porté sur un circuit appelé LDPC, dont l'acronyme signifie « Low Density Parity Check » ou « Vérificateur de parité à faible densité » [Gal'62]. C'est un code correcteur d'erreur linéaire utilisé pour transmettre un message à travers un canal de transmission bruité.

En l'absence de code correcteur, le bruit additionné au signal émis peut introduire des erreurs pseudo-aléatoires, qui ne peuvent alors pas être détectées par le récepteur et encore moins être corrigées (voir l'illustration figure ).

Mots à coder (3 bits) $b$ de données	Matrice de génération $G$	Mots codés (5 bits) $w=b \times G$	Matrice de vérification de parité $H$
(1 1 1)	1 0 0 0 1	(1 1 1 0 1)	1 0 0 1 1
(1 0 0)	0 1 0 1 1	(1 0 0 0 1)	0 1 1 1 0
(0 1 0)	0 0 1 1 1	(0 1 0 1 1)	
$(b_0 b_1 b_2)$		$(b_0 b_1 b_2 c_0 c_1)$	
Décodage	Si $w \times H^T = (0 \ 0)$ alors $(b_0 b_1 b_2)$ ok Sinon correction :		
(0 1) pb sur $(b_1 \text{ ou } b_2)$ corriger avec $b_0 c_0 c_1$	Mots reçus $w(b,c)$ de 5 bits		
(1 0) pb sur $(b_0 \text{ ou } c_1)$ corriger avec $b_1 b_2 c_0$	3 bits $b$ de données et <b>2 bits <math>c</math> de contrôle</b>		
(1 1) pb sur $(c_0 \text{ ou } c_1)$ (erreur sur les bits de contrôle)	(1 1 1 0 1) $\Rightarrow$ (0 0) ok		
	(1 0 0 1 1) $\Rightarrow$ (1 1) pas de correction		
	(1 1 0 1 1) $\Rightarrow$ (1 0) $b_0$ à corriger car $c_0=1$		

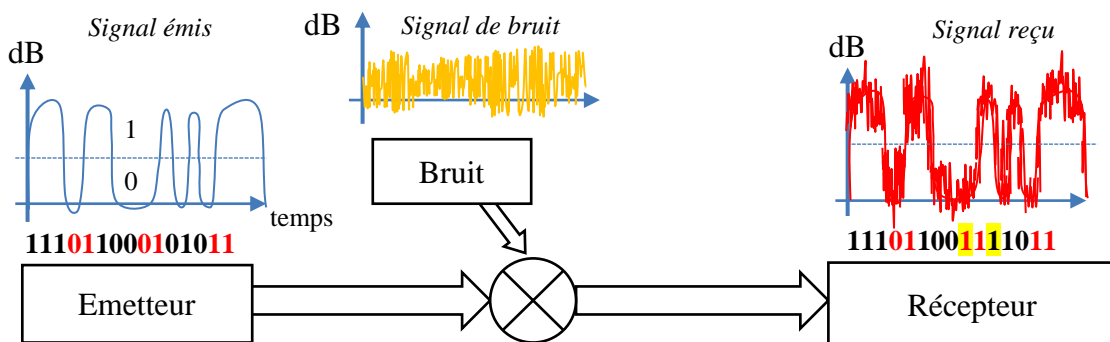


figure : Principe de l'algorithme de codage / décodage intégré dans le LDPC

Le code a pour rôle d'introduire de la redondance dans le signal émis afin de pouvoir réaliser une correction si un ou plusieurs bits sont mal transmis. Pour cela, les bits de données sont associés à des bits de contrôle obtenus à l'aide d'une matrice de codage générant un mot binaire plus long en nombre de bits mais qui, une fois propagé dans le canal bruité et

intercepté par le récepteur, peut être décodé et corrigé, assurant une transmission du message plus fiable et robuste. Cet algorithme, proposé pour la première fois par Robert Gallager, est basé initialement sur le théorème de Shannon. Il est codé à l'aide de la matrice de génération  $G$  et décodé à l'aide de la matrice de la parité  $H$  : ces matrices sont liées mathématiquement. Ainsi comme dans l'exemple précédant (figure ), un mot codé sur 3 bits devient un mot codé sur 5 bits (3 bits de données et 2 bits de contrôle). Les bits de données ne changent pas puisque la matrice  $G$  est basée sur la matrice identité, par contre les bits de contrôle sont obtenus par un calcul matriciel et dépendent des bits de données. En fonction de la construction du code, il est plus ou moins facile de corriger et de retrouver le mot initial même si un bit est erroné. A titre d'exemple, le code simplifié présenté figure , n'est ni optimal ni complet, il permet juste une compréhension simple et claire du codage du LDPC.

Le principe évoqué ci-dessus est intégré dans le circuit LDPC, d'une manière beaucoup plus complexe et performante afin d'être compatible avec le standard IEEE, 802.11n [I3e'09]. Ce standard est utilisé dans les applications Wifi. Le bloc LDPC est également implémenté dans d'autres applications comme la transmission de signaux pour la télévision numérique « DVB-S2 » [Ets'09].

Chez STMicroelectronics, cette architecture de LDPC, incluant plusieurs taux de code configurable, a été utilisée comme circuit de test de référence pour évaluer différents nœuds technologiques, à savoir en BULK 0,13 $\mu$ m [Ura'05], 65nm [Ura'08] ou 45nm [Mou'10]. Il s'agit donc un circuit de test bien connu, pour lequel il existe déjà des mesures silicium permettant de faire des comparaisons avec les résultats obtenus au cours de cette thèse.

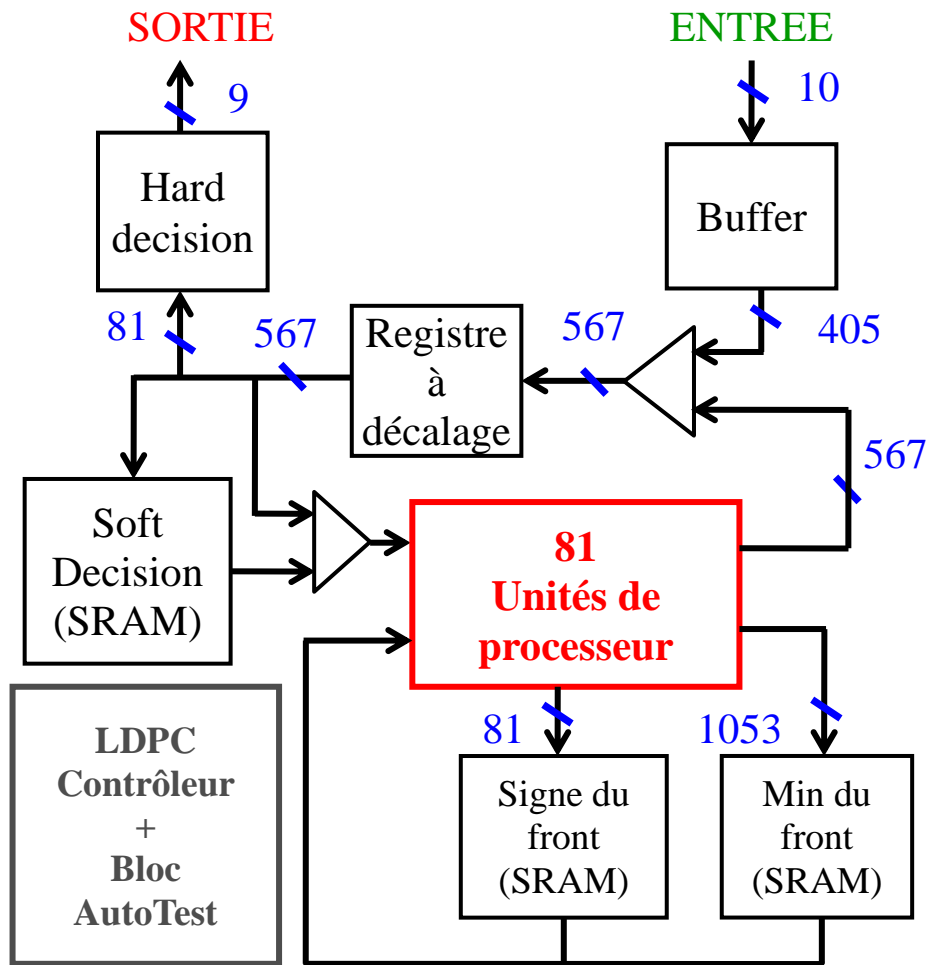


figure : Architecture du bloc logique LDPC

L'architecture du bloc LDPC sélectionné (figure ) est principalement composée d'un codeur, d'un décodeur, de mémoires et d'un bloc d'autotest. Le bloc autotest permet de réaliser en interne une vérification fonctionnelle du bloc en injectant dans le canal de transmission un bruit virtuel engendrant des erreurs. Comme ces erreurs sont connues, l'objectif de ce bloc est de vérifier que l'architecture du LDPC est bien capable de corriger un maximum de ces erreurs en fonction du taux d'erreur binaire appelé BER, pour « Bit Error Rate » en anglais. Ce taux est relatif et mesuré à la réception du signal, il est dû aux perturbations rencontrées sur le canal de transmission. Il est égal au nombre de bits erronés divisé par le nombre total de bits transmis.

Le bloc logique LDPC a tout d'abord été synthétisé et placé-routé en utilisant la plateforme de conception de la technologie 65nm LP BULK. Des interrupteurs de puissance de la technique MTCMOS ont été ajoutés pour réduire la consommation statique de ce bloc et servir de point de comparaison.

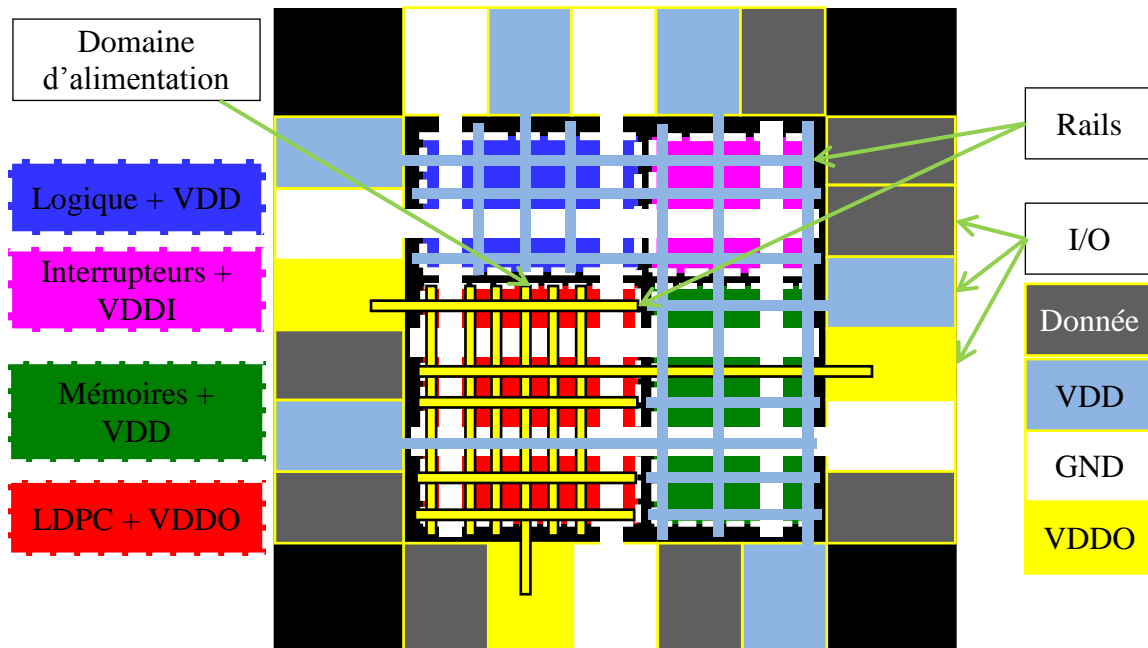
Ce circuit a ensuite été porté en 65nm LP PD-SOI après avoir été validé en termes de contraintes temporelles. Dans ce cas, et par construction, tous les transistors sont de type FB, y compris les interrupteurs de puissance. C'est pour cela que nous avons étudié deux autres versions avec des interrupteurs de puissance issus de la plateforme de conception PD-SOI : la première inclut des transistors de puissance de type BC, fonctionnellement équivalents aux interrupteurs de puissance implémentés en BULK afin de déterminer le gain intrinsèque apporté par la technologie PD-SOI ; la deuxième version inclut la solution Auto-DTMOS développée dans le chapitre précédent, permettant d'estimer le gain supplémentaire apporté par cette technique circuit.

### 4.2. Stratégie d'insertion des interrupteurs de puissance pour circuit LDPC

Dans ce paragraphe, la stratégie d'insertion des interrupteurs de puissance unitaires et la méthodologie de dimensionnement du réseau ainsi formé sont introduits.

#### 4.2.1. Conception d'un circuit

Dans le flot de conception numérique, une fois le LDPC synthétisé à l'aide des cellules standards, il faut passer à l'étape qui consiste à placer ces cellules standards et à les router. La première sous-étape du « Back-End » définit le plan du circuit où seront implémentés tous les blocs. Cette sous-étape peut s'appeler le « floorplan ». On y insère aussi les différents domaines d'alimentation, mais également les futures connections des signaux d'entrée/sortie, ainsi que les rails d'alimentation et de masse.



**figure : Exemple de « floorplan » incluant un domaine d'alimentation VDDO**

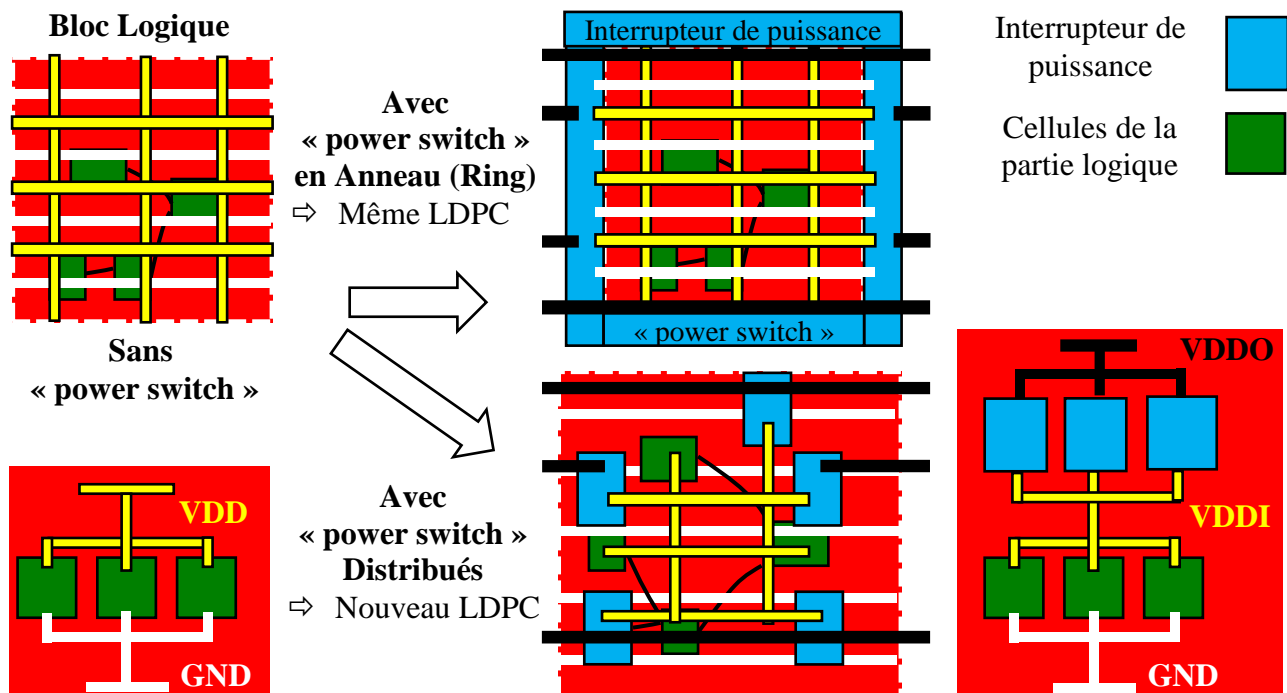
L'insertion d'un domaine d'alimentation séparé VDDO (figure ) va permettre l'insertion d'un bloc LDPC alimenté de manière indépendante avec ou sans interrupteurs de puissance. Concernant l'autre domaine d'alimentation global, VDD, il permet l'alimentation du reste du circuit comme notamment la logique de contrôle, les différentes mémoires et les blocs analogiques.... Un domaine d'alimentation peut être créé grâce à l'insertion d'interrupteurs de puissance qui viennent séparer l'alimentation global VDD de l'alimentation virtuelle, VDDI. Une fois cette sous étape réalisée, il faut placer et router toutes les cellules :

- les cellules d'entrée/sortie ou IO (Alimentation, Donnée...)
- les cellules standards composant les blocs logiques (LDPC...)
- les mémoires (SRAM, ROM)
- les blocs analogiques (capteurs, band gap)
- la bibliothèque MTCMOS dans les zones préalablement dédiées.

#### 4.2.2. Insertion des interrupteurs de puissance en anneau

L'insertion des interrupteurs de puissance se fait en anneau autour du domaine d'alimentation qu'ils doivent couper ou alimenter. L'implémentation en anneau permet de conserver au centre de la zone dédiée, un espace libre permettant le placement et le routage de la partie logique. Les outils de conception permettent aussi d'implémenter les interrupteurs de

puissance de manière distribuée et mélangée aux cellules élémentaires de la partie logique, mais nous n'avons pas pris cette option.



**figure : Impact sur le bloc logique d'une implémentation de bibliothèque MTCMOS avec interrupteurs de puissance en anneau ou distribués**

Dans notre cas, l'implémentation en anneau permet de conserver un placement et routage intact de la partie logique, qu'elle soit associée ou non à un réseau d'interrupteurs de puissance. Avec une implémentation distribuée, il est nécessaire de refaire un placement et routage (figure ) afin d'aménager l'espace pour y intégrer de manière homogène les interrupteurs de puissance unitaires. Le placement des interrupteurs de puissance de manière distribuée peut être avantageux en termes de chute de tension, de distribution équilibrée du potentiel VDDI et de densité d'intégration, car les transistors de puissance peuvent remplacer, dans ce cas, les « filler » ou autres zone libres induits par le routage. Cependant, pour des raisons de flexibilité et dans le but de porter directement le circuit LDPC de la technologie BULK vers la technologie PD-SOI sans à avoir à refaire de synthèse ni de placement/routage, le choix a été fait d'implémenter le réseau d'interrupteurs de puissance en anneau. De cette manière, il est possible de dissocier la partie logique des interrupteurs de puissance. Les deux parties pouvant être conçues indépendamment dans des flots de conception différents.

## 4.2.3. Dimensionnement des interrupteurs de puissance

Le nombre de transistors unitaires,  $N_{SW}$ , à implémenter dépend de la partie logique à laquelle le réseau d'interrupteurs est associé. Une première étude consiste à connaître les différents courants de la partie logique, le courant d'appel (courant maximal et courant dynamique moyen) ainsi que le courant statique. Cela permet de connaître la résistance maximale équivalente du réseau d'interrupteurs de puissance admissible  $R_{on}$  [Mut'99], mais aussi d'évaluer rapidement le gain en termes de réduction de la consommation statique et donc l'opportunité d'utiliser des transistors de puissance. En effet, si le courant de veille de la partie logique  $I_{ddq}$  n'est pas supérieur de plus d'une décade au courant en mode OFF des interrupteurs de puissance, il n'est pas vraiment nécessaire d'insérer des interrupteurs. En LP 65nm PD-SOI, l'implémentation du bloc logique LDPC FB LVt induit un courant de veille  $I_{ddq}$  supérieur entre 3 et 4 décades au courant de fuite  $I_{off}$  des interrupteurs de puissance BC HVt.

La résistance du réseau d'interrupteurs de puissance,  $R_{on}$ , est proportionnelle à l'inverse du nombre d'interrupteurs unitaires  $N_{SW}$  mis en parallèle et abutés. Cette résistance engendre une chute de tension,  $V_{drop}$ , qui est fonction du courant dynamique consommé par la partie logique. La spécification en termes de fréquence de fonctionnement du circuit alimenté détermine la chute de tension maximale autorisée et ainsi le nombre d'interrupteurs de puissance à utiliser. Plus ce nombre est grand, plus la chute de tension  $V_{drop}$  sera faible.

La valeur de la résistance  $R_{on_{unitaire}}$  d'un transistor de puissance dépend de la température, de la tension d'alimentation et de qualité du procédé de fabrication. Il convient alors de garantir que la valeur de cette résistance respecte les spécifications imposées par le cahier des charges quelle que soit la configuration dans laquelle le circuit peut se trouver. En l'occurrence, trois cas vont intervenir : le cas typique, le meilleur cas et le cas critique. Le cas typique est défini pour un circuit soumis à des conditions d'alimentation, de température et de procédé nominales. Le pire cas est défini pour une tension d'alimentation faible (10% inférieure à la tension nominale), une température élevée (car pas de phénomène d'inversion en température pour cette technologie à cette tension) et un procédé de fabrication défavorable. C'est dans ce cas que la résistance série de l'interrupteur de puissance est la plus élevée, mais aussi que le courant d'appel maximal du circuit alimenté est le plus faible. Enfin, le meilleur cas est défini pour une tension d'alimentation élevée (10% supérieure à la tension nominale), une température faible (idem : pas de phénomène d'inversion en température pour cette technologie à cette tension) et un procédé favorable. Dans ce cas, en raison d'un appel

en courant important la chute de tension «  $V_{drop}$  » risque d'être la plus élevée, entraînant une dégradation relative de délai importante. Mais comme la fréquence de fonctionnement possible est bien plus grande lorsque la tension d'alimentation est élevée, elle respecte toujours la fréquence d'horloge spécifiée par le cahier des charges : il n'y a pas d'intérêt à s'attarder sur ce point de fonctionnement « meilleur cas ».

Dans le tableau , un nombre d'interrupteurs de puissance unitaires NSW est instancié pour le réseau de type BC et Auto-DTMOS pour alimenter le LDPC. Afin d'atteindre la même valeur de résistance  $R_{on}$  et donc la même chute de tension maximale  $V_{drop}$  de 10mV, dans le premier cas (A), le nombre NSW est de 3675 pour le BC et de 2887 pour la solution Auto-DTMOS à  $V_{DD}=1,2V$ ,  $T=25^{\circ}C$  et pour un procédé typique ; alors que dans le deuxième cas (B), le nombre NSW est respectivement de 4394 et de 3511 pour le point de fonctionnement le plus pessimiste, à savoir  $V_{DD}=0,8V$ ,  $T=125^{\circ}C$  et un procédé défavorable.

<b>Point de fonctionnement</b> <b>Caractéristiques du</b> <b>transistor de puissance</b> <b>unitaire</b>	<b><math>V_{DD}=1,2V</math></b> <b><math>T=25^{\circ}C</math></b> <b>Procédé</b> <b>typique</b>	<b><math>V_{DD}=0,8V</math></b> <b><math>T=25^{\circ}C</math></b> <b>Procédé</b> <b>typique</b>	<b><math>V_{DD}=0,8V</math></b> <b><math>T=125^{\circ}C</math></b> <b>Procédé</b> <b>défavorable</b>
Ron BULK et BC PD-SOI ( $\Omega$ )	122,5	308	325,5
Ron Auto-DTMOS PD-SOI ( $\Omega$ )	96,25	210	260,1
Imax (LDPC) (mA)	300	150	135
(A) $V_{drop}$ (mV) : BC   Auto-DTMOS	10	12,57   10,91	11,95   12,16
(B) $V_{drop}$ (mV) : BC   Auto-DTMOS	8,36   8,22	10,51   8,97	10

**tableau : Comparaison des chutes de tensions en fonction du point de fonctionnement du circuit et du nombre d'interrupteur de puissance NSW (A) et (B).**

Le tableau est obtenu par simulation sans tenir compte des éventuelles résistances, capacités et inductances parasites des entrées/sorties et des interconnexions. Le courant maximal du LDPC,  $I_{max}$ , est issu d'une estimation donnée par les outils de conception. Que ce soit en typique ou dans un cas pessimiste, le gain en  $R_{on}$  de la solution BC par rapport à la solution Auto-DTMOS est approximativement de -20%. A  $V_{DD}=0,8V$ ,  $T=25^{\circ}C$  et pour un procédé typique, la polarisation FBB plus élevée induit une meilleure réduction de  $R_{on_{unitaire}}$ . Cependant, ce point de fonctionnement n'est pas assez pessimiste ; il permet juste de mettre en avant le fait que la solution Auto-DTMOS est meilleure, induisant une plus faible chute de tension  $V_{drop}$  et donc impactant moins la vitesse du LDPC.



Le chiffre du courant d'appel du LDPC estimé et maximal,  $I_{max}$ , utilisé pour calculer  $R_{on}$ , garantit une certaine fiabilité en introduisant du pessimisme dans le dimensionnement et dans la définition du nombre NSW d'interrupteurs de puissance à implémenter. Le courant estimé est celui d'un bloc LDPC synthétisé en LP 65nm BULK atteignant une fréquence de 200 MHz pour une tension de 1,05V, une température de 125°C et un procédé lent. En réalité la tension recommandée pour alimenter ce type de circuit en technologie LP 65nm BULK et PD-SOI est au minimum de 1,05V et au maximum de 1,35V, centrée autour d'une alimentation nominale de 1,2V. Dans le cas du démonstrateur proposé pour valider et comparer les différentes solutions d'interrupteurs de puissance associés au LDPC en BULK et PD-SOI, la tension d'alimentation du circuit balayera une valeur allant de 1V à 1,4V afin d'élargir le spectre des conditions d'utilisation.

Pour le pire cas, plus de 4394 interrupteurs de puissance sont recommandés pour la solution BC et plus de 3511 pour la solution Auto-DMTOS. En considérant une marge d'environ 10%, le nombre NSW final qui va constituer le réseau en forme de ring et qui va alimenter le LDPC est de 4800 pour les solutions BULK et BC pour la technologie PD-SOI et de 3840 pour la solution spécifique à la technologie PD-SOI Auto-DMTOS. Ce chiffre permet de respecter dans le pire cas une chute de tension maximale de 10mV pour un courant d'appel du LDPC pessimiste et pour une tension minimale VDD de 1V. Nous conservons la réduction de 20% du nombre d'interrupteurs de puissance, NSW, de la solution BC vers la solution Auto-DMTOS, pour le même  $R_{on}$ . Dans le cas de la solution Auto-DMTOS, la surface gagnée est de 20%. Cette surface libérée a été avantageusement exploitée pour insérer des capacités de découplage. Comme il a été convenu d'insérer un bloc de polarisation tous les 1000 interrupteurs de puissance, le nombre de blocs de polarisation auto-adaptative pour 3840 interrupteurs de puissance est de 4. Ce chiffre permet, en plus, de répartir de manière équilibrée les blocs de polarisation : un sur chaque côté du réseau qui est implémenté en anneau. Il correspond à un bloc de polarisation tous les 960 interrupteurs de puissance unitaires.

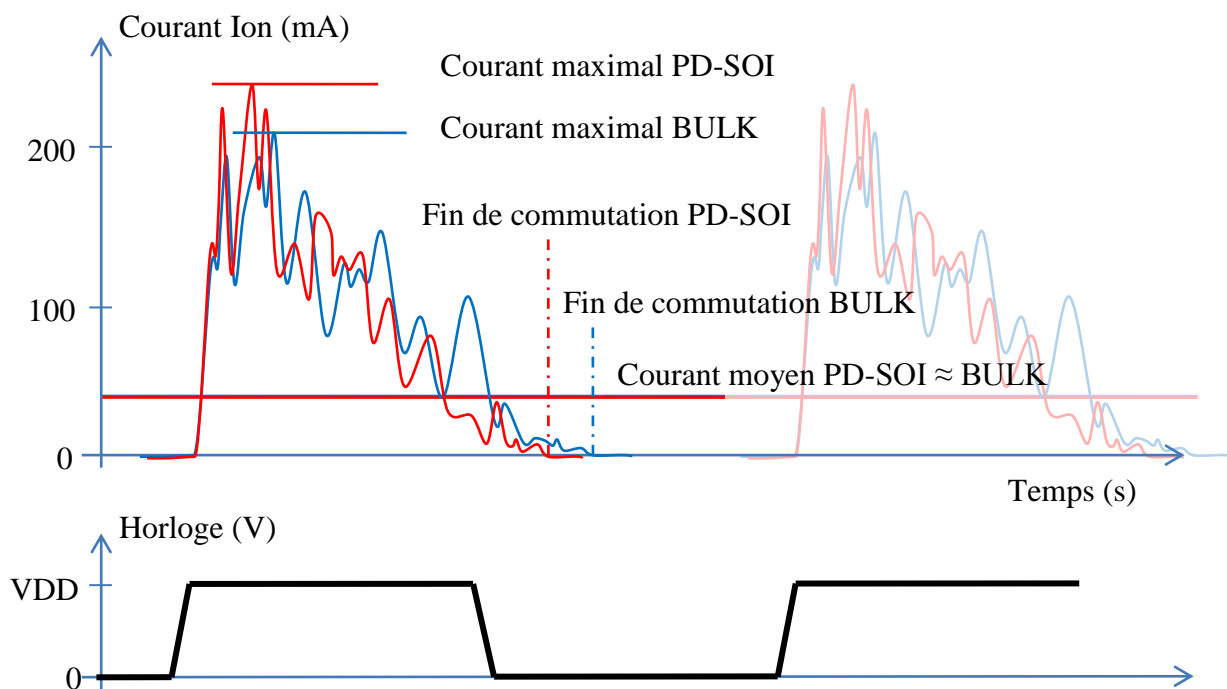
### 4.3. Comparaison des réalisations silicium BULK et PD-SOI

#### 4.3.1. Nombre d'interrupteurs de puissance unitaires NSW

Les interrupteurs de puissance unitaires de type BULK et BC PD-SOI ont les mêmes résistances unitaires  $R_{on_{unitaire}}$  et  $R_{off_{unitaire}}$ . Le transistor de puissance PD-SOI Auto-DMTOS

a une résistance  $R_{on\_unitaire}$  globalement 20% plus faible que les deux autres pour la même résistance  $R_{off\_unitaire}$  puisqu'ils se retrouvent tous avec la même polarisation de body,  $V_{bs}=0$ , en mode OFF. Cela signifie que le réseau en anneau, constitué de 4800 unités pour la solution BULK et BC en PD-SOI aura la même résistance  $R_{on}$  totale que la solution Auto-DTMOS constituée de 3840 « power switch » unitaires. La résistance totale  $R_{off}$  du réseau Auto-DTMOS est en revanche 25% plus élevée que celle des deux premiers réseaux, due au nombre NSW réduit de 20%. Cela permet de réduire encore le courant  $I_{off}$ , et la surface d'implémentation, des interrupteurs de puissance de 20%.

Par contre, les parties logiques BULK et PD-SOI (constituée de MOS FB), auront des courants d'appel bien différents. En effet, le courant du LDPC dépend principalement du type de MOS qu'il contient. Qu'il soit issu d'un portage direct ou entièrement conçu dans une plateforme PD-SOI, le courant d'appel des LDPC en PD-SOI est sensiblement plus élevé qu'en BULK. Car les transistors de type FB sont capables de délivrer plus de courant, ce qui permet d'augmenter de manière significative la vitesse, puisqu'ils assurent une commutation plus rapide. A même tension d'alimentation,  $V_{DD}$ , le courant maximal d'appel du LDPC en technologie PD-SOI est plus élevé qu'en technologie BULK pour un courant moyen quasi identique. La signature en courant est différente, le courant d'appel est plus élevé mais la fenêtre du courant d'appel est plus petite en PD-SOI qu'en BULK (figure ).



**figure : Exemple de signature en courant du LDPC BULK et PD-SOI pour la même tension d'alimentation VDD**

L'intégrale du courant en fonction du temps, qu'il soit en BULK ou en PD-SOI, donne approximativement le même résultat. Il représente la quantité de charge, et donc la capacité totale à faire commuter, à VDD constant. Cela démontre que la capacité totale à faire commuter est sensiblement la même d'une technologie à l'autre et que la technologie PD-SOI est capable de les faire commuter plus rapidement. Le courant moyen reflète assez bien ceci en donnant une idée de la capacité totale à faire commuter à chaque coup d'horloge.

En technologies LP 65nm BULK et PD-SOI, le nombre de niveaux métalliques de routage est de 7. Ce chiffre aide les outils de routage en leur offrant plus de libertés pour router les différentes cellules ; mais il induit une résistivité et une capacité des interconnexions de plus en plus élevées. En s'appuyant sur le fait que la capacité des interconnexions des LDPC en technologies BULK et en PD-SOI est la même à cause du portage direct et que les capacités intrinsèques du transistor FB PD-SOI sont plus faibles que celles du MOS BULK, il est évident que la capacité totale est majoritairement liée à la capacité des interconnexions puisqu'il est quasiment impossible de distinguer les deux technologies en matière de capacité de circuit totale à commuter.

Pour résumer, en technologie PD-SOI, le MOS FB est capable de fournir plus de courant que son homologue BULK, ce qui lui permet de faire commuter plus rapidement toutes les capacités de son circuit. Comme les capacités des interconnexions sont identiques à cause du portage direct, il exécute cette commutation dans une fenêtre temporelle plus petite et ceci pour une même tension d'alimentation VDD. En technologie PD-SOI, le pic de courant est plus élevé mais il a une consommation dynamique moyenne par cycle d'horloge équivalente au BULK. La différence entre les technologies BULK et PD-SOI vient du fait que la fréquence du LDPC peut être plus élevée en PD-SOI pour la même tension d'alimentation. Pour la même résistance du réseau d'interrupteurs de puissance  $R_{on}$ , le pic en courant étant supérieur en PD-SOI, il peut légèrement accroître la chute de tension aux bornes de  $R_{on}$ ,  $V_{drop}$ , ainsi que la dégradation relative de délai. Cela n'empêche en rien le respect du cahier des charges car, même si la dégradation de délai relative est supérieure en PD-SOI, passant de 3% en BULK à 3,5% en PD-SOI à titre d'exemple, la fréquence opérationnelle reste bien supérieure en PD-SOI, avec un gain toujours autour de +20% pour une tension d'alimentation VDD nominale et identique.

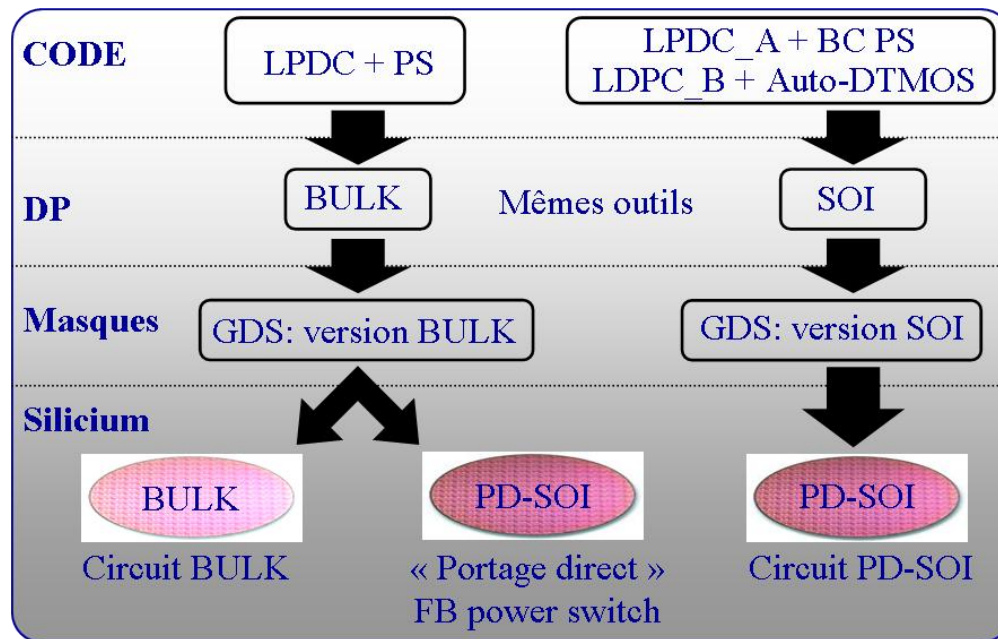
Plusieurs circuits vont permettre de valider la technologie et la plateforme PD-SOI en y incluant la solution MTCMOS proposée « Auto-DTMOS » et ainsi faire une comparaison

avec la solution de référence BC sans polarisation de body en PD-SOI et la solution en technologie BULK.

Un premier circuit contenant un bloc LDPC est conçu et fabriqué en BULK. Le bloc LDPC est alimenté via un réseau d'interrupteurs de puissance classique afin de réduire sa consommation statique. Ce circuit va être porté en technologie PD-SOI. Il permet de valider le portage direct d'un circuit entièrement conçu avec une plateforme de conception BULK et fabriqué sur une tranche de silicium PD-SOI. La mesure de la fréquence maximale et de la consommation dynamique du bloc LDPC a pour objectif de présenter le gain technologique du BULK vers le PD-SOI. Comme par construction, tous les MOS sont de type FB, y compris les interrupteurs de puissance, ce circuit n'est pas optimisé en PD-SOI en termes de consommation statique.

Un second circuit entièrement conçu et fabriqué en PD-SOI est proposé incluant deux blocs LDPC. Le premier bloc LDPC est issu du portage direct (LDPC\_A) et le second bloc LDPC a été synthétisé en PD-SOI (LDPC\_B) avec une fréquence 20% plus élevée. Comme pour le premier circuit, chaque LDPC est alimenté via un réseau d'interrupteurs de puissance indépendant, PS : le premier bloc LDPC, avec la solution BC, le second avec la solution « Auto-DTMOS » et 20% en moins d'interrupteurs de puissance. Ces circuits optimisés en technologie PD-SOI auront pour objectif de montrer, en plus du gain technologique, un gain lié à la plateforme de conception PD-SOI, notamment en termes de réduction de la consommation statique.

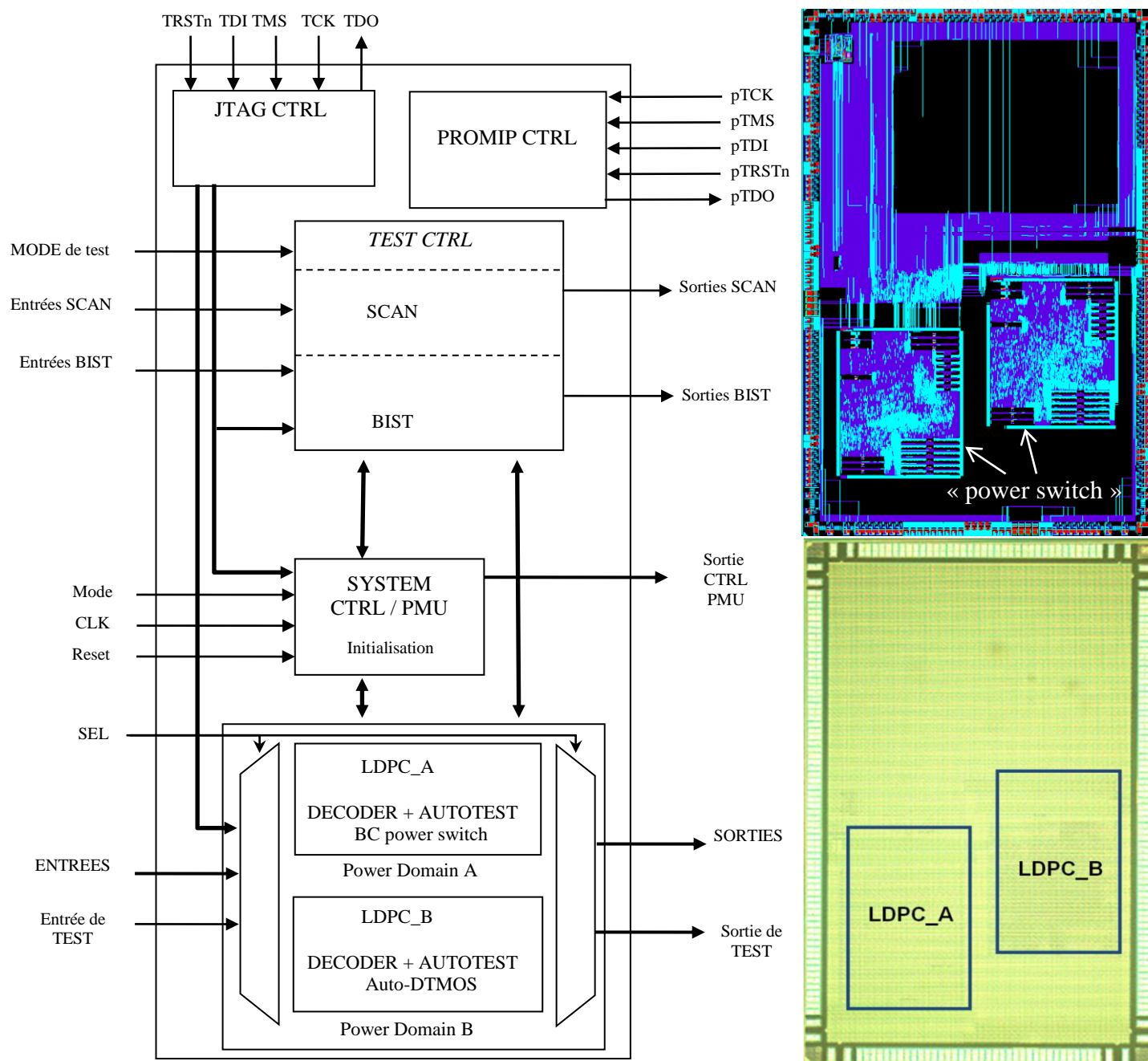
La figure présente les différents circuits qui seront disponibles et comparés.



**figure : Flots de conception des circuits démonstrateurs**

Le premier circuit en technologie BULK contient un LDPC mais également d'autres blocs comme les IO, la bibliothèque MTCMOS, les mémoires mais aussi l'unité de gestion de puissance « PMU », les blocs de programmation et de test internes comme le sont respectivement le « JTAG » [I3e'90] et le « BIST ». La gestion des états du réseau d'interrupteurs de puissance se fait par le contrôleur mais comme celui n'est pas directement accessible par l'extérieur, l'unité de gestion de puissance s'en occupe. Son rôle étant aussi de gérer l'arbre d'horloge et bien d'autres modes.

En BULK, le circuit contient un domaine d'alimentation avec un LDPC de type LVt alimenté par un réseau d'interrupteurs de puissance classique, HVt. Ce circuit porté en PD-SOI sera entièrement composé de MOS de type FB ; l'interrupteur de puissance est alors FB et fortement fuyant. En PD-SOI le circuit contient deux domaines d'alimentation, un pour chaque LDPC, alimentés via deux réseaux indépendants d'interrupteurs de puissance peu fuyants BC et Auto-DTMOS (figure ).



**figure : Architecture du circuit PD-SOI avec vue layout et photographie du circuit**

Les circuits en technologie BULK et PD-SOI ont la même dimension ( $8\text{mm}^2$ ) et le même nombre d'entrées/sorties (224). La taille du LDPC est la même pour tous ( $1,5\text{mm}^2$ ) et ils ont été implémentés avec des transistors de type FB en PD-SOI. Les interrupteurs de puissance utilisés sont de type HVt. Le reste des blocs utilisés sont de type SVt et HVt.

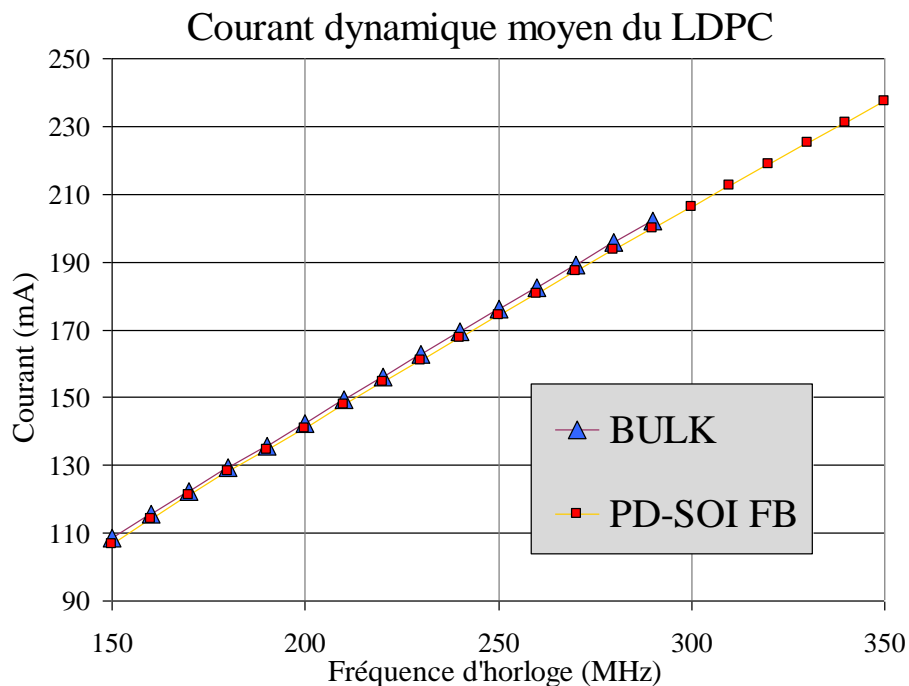
## 4.3.2. Vitesse et puissance dynamique mesurées du LDPC

Le portage direct du LDPC conçu en BULK à 200MHz vers la technologie PD-SOI montre très peu de différence en termes de portes logiques et de consommation par rapport à un LDPC entièrement conçu en PD-SOI avec pour une fréquence de 240 MHz. Le LDPC issu du portage direct en PD-SOI est en effet fonctionnel pour une fréquence maximale 20% plus élevée qu'en BULK soit 240MHz. Il est donc aujourd'hui la meilleure solution en termes de conception numérique car il permet un gain de temps élevé en réutilisant des blocs souvent déjà existants et en considérant une fréquence d'horloge 20% plus élevée. La seule étape de vérification est celle des contraintes de temps : vérifier que le circuit est fonctionnel pour une fréquence maximale spécifiée et qu'il ne risque pas de rencontrer des violations de maintien.

En passant du BULK vers le PD-SOI, le gain technologique en vitesse attendu pour un même circuit est aux alentours de +20% en typique ( $V_{DD}=1,2V$ ,  $T=25^{\circ}C$  et un procédé nominal). Il est également possible de faire fonctionner le circuit à la même fréquence d'horloge en réduisant la tension d'alimentation  $V_{DD}$  aux alentours de -15%. En réduisant  $V_{DD}$ , le courant des MOS FB devient plus faible, de même que la quantité de charges à faire commuter ; la consommation dynamique est donc réduite, c'est le second gain technologique. Comme le MOS FB est capable de délivrer plus de courant que son homologue BULK, il permet d'atteindre la même vitesse à plus faible tension d'alimentation. De plus, le courant maximal d'appel devient plus faible ainsi que la chute de tension,  $V_{drop}$ . Cela permet de retrouver en BULK et en PD-SOI quasiment la même dégradation relative de délai, voire moins, en introduisant la même résistance du réseau d'interrupteurs de puissance  $R_{on}$  pour alimenter le LDPC.

Tous les circuits qui ont été fabriqués sont fonctionnels. La première comparaison entre la technologie BULK et la technologie PD-SOI est faite sur le bloc logique LDPC lorsque les interrupteurs de puissance sont passants. Tous les LDPC en technologie PD-SOI sont quasiment similaires qu'il soit issu du portage direct, et donc synthétisé en BULK à 200MHz, ou qu'il soit synthétisé en PD-SOI à une fréquence de 240MHz. Ils ont à peu près le même nombre de portes logiques, la même fréquence de fonctionnement, et la même consommation dynamique pour la même surface. On ne distinguera donc que 2 LDPC, le premier en BULK et le second en PD-SOI, qu'il soit porté à partir BULK vers le PD-SOI ou directement synthétisé en PD-SOI.

La figure montre la courbe de courant dynamique moyen en fonction de la fréquence d'horloge en typique des circuits LDPC BULK et PD-SOI.



**figure : Courant dynamique moyen mesuré sur silicium des deux LDPC, BULK et PD-SOI en fonction de la fréquence d'horloge**

La consommation dynamique à fréquence égale est sensiblement la même en BULK et en PD-SOI, puisqu'elle dépend majoritairement des résistances et capacités équivalentes des interconnexions qui sont les mêmes en BULK et en PD-SOI. Avec la technologie PD-SOI, il est possible d'atteindre une fréquence fonctionnelle moyenne de 350MHz dans des conditions quasiment nominales, alors qu'en BULK, elle est au maximum à 290MHz. Cela montre bien que la technologie PD-SOI a des transistors qui permettent de délivrer plus de courant que ceux en BULK, ce qui permet de faire commuter plus rapidement le circuit. En synthétisant en PD-SOI à la même fréquence (200MHz) que le BULK, la configuration de la connexion des portes ainsi que leur nombre sont différents. Le chemin critique peut avoir une profondeur logique supérieure : cela permet de sérialiser la partie logique afin de faire commuter moins de portes logiques et d'obtenir une courbe de puissance dynamique par fréquence plus faible. L'avantage du portage direct provient principalement du fait que la fréquence maximale étant plus élevée à même tension d'alimentation VDD (tableau ), il est possible de réduire la tension d'alimentation de quelques pourcents en PD-SOI (tableau ).



Fréquence maximale & Puissance dynamique par MHz			
VDD=1,2V	BULK 65LP	PD-SOI 65LP	PD-SOI vs BULK
Fréquence (MHz)	300	360	+ 20 %
Puissance (mW/MHz)	0,81	0.8	- 1 %

**tableau : Comparaison à même tension d'alimentation de la fréquence maximale et de la puissance dynamique par MHz en BULK et PD-SOI**

Le tableau montre qu'à même tension d'alimentation VDD=1,2V, le LDPC en PD-SOI a une fréquence maximale 20% plus élevée pour la même puissance dynamique par MHz. Ce résultat démontre bien que le circuit en PD-SOI porté à partir du BULK est plus rapide et que la consommation dynamique par MHz dépend principalement des interconnexions et de l'alimentation VDD.

Puissance dynamique @ même fréquence			
Fréquence = 360MHz	BULK 65LP	PD-SOI 65LP	PD-SOI vs BULK
VDD (V)	1,4	1,2	- 14 %
Puissance (mW)	406	288	- 30 %

**tableau : Comparaison à même fréquence maximale de la tension d'alimentation et de la puissance dynamique en BULK et PD-SOI**

Le tableau présente la tension d'alimentation VDD qui permet d'atteindre la même fréquence maximale de 360MHz. En BULK, il faut augmenter la tension d'alimentation de 200mV, pour atteindre dans ce cas 1,4V. La tension d'alimentation 14% plus faible en technologie PD-SOI permet de réduire la consommation dynamique de 30%.

En conclusion, la technologie 65nm LP PD-SOI permet de travailler soit avec une fréquence moyenne 20% plus élevée qu'en BULK soit à même fréquence avec une consommation dynamique 30% plus faible. Ces deux gains technologiques ne peuvent être exploités qu'à la condition, que la consommation statique du PD-SOI soit identique à celle de la technologie BULK.

## 4.3.3. Puissances statiques mesurées

## 4.3.3.1. Interrupteurs de puissance passants : LDPC alimenté

La puissance statique étant le point faible de la technologie PD-SOI, une seconde comparaison est faite, toujours, lorsque les interrupteurs de puissance sont passants mais lorsque la technique de « clock gating », ou isolation d'horloge, désactive le signal d'horloge entraînant l'interruption de la commutation du bloc LDPC. Dans ce cas, le courant de fuite appelé  $I_{ddq}$ , est le courant de fuite de la partie logique inactive.

Puisque le type de MOS utilisé est de type FB en PD-SOI, Le LDPC de technologie PD-SOI (3,4mW) consomme 6 fois plus qu'en BULK (560 $\mu$ W) allant jusqu'à 10 fois plus pour des points de fonctionnements critiques,  $V_{DD}=1,35V$  notamment. Ce sont ces chiffres ainsi que ce ratio qui apparaîtraient si aucune technique de la réduction de la consommation n'était utilisée à même tension d'alimentation.

Puissance statique @ même fréquence maximale ciblée			
LDPC alimenté	BULK 65LP	PD-SOI 65LP	PD-SOI vs BULK
VDD (V)	1,4	1,2	- 14 %
Puissance statique (mW)	1,2	3,4	<b>x 2.8</b>

**tableau : Comparaison à même fréquence maximale ciblée de la puissance statique du bloc LDPC alimenté en BULK et PD-SOI**

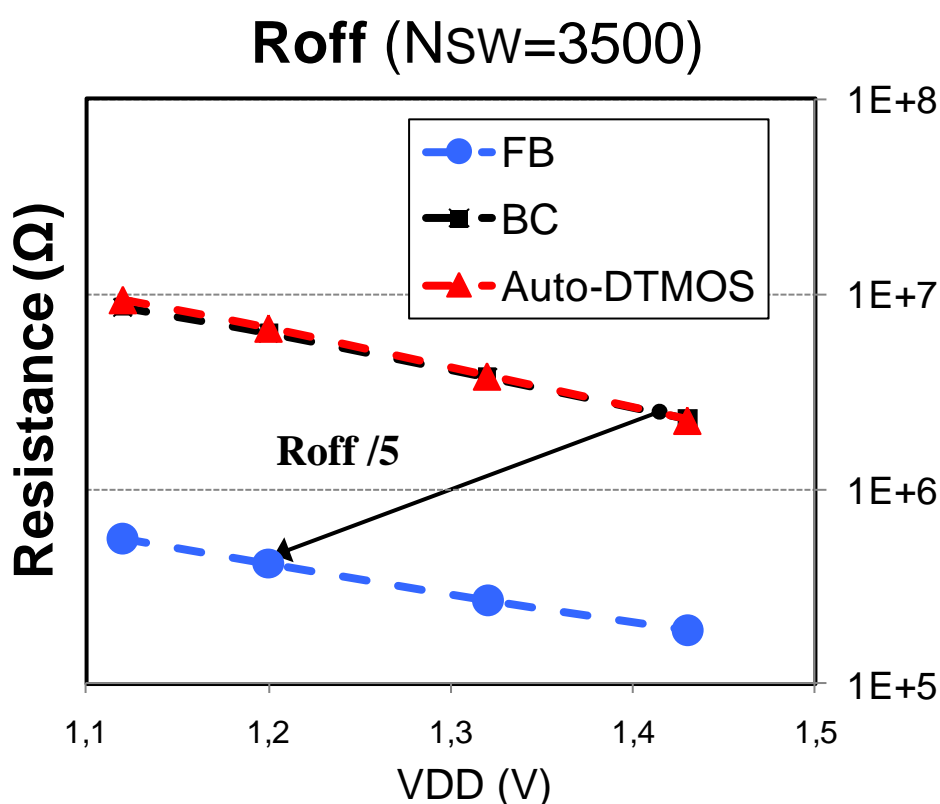
Si l'on compare cette fois à même fréquence de fonctionnement (360MHz), le ratio de la consommation statique entre la technologie PD-SOI et la technologie BULK est réduit de 6 à 2,8 (tableau ). Alors que ce ratio, pour un circuit microprocesseur, est de 55 dans [Pot'09]. La consommation statique en PD-SOI reste plus importante qu'en BULK en raison de l'effet de body flottant des transistors FB qui compose la partie logique malgré une diminution de 14% de la tension d'alimentation. Ce chiffre montre malgré tout, que la consommation

statique d'un circuit PD-SOI sans technique de réduction de la consommation statique reste supérieure à celle d'un circuit en technologie BULK.

#### 4.3.3.2. Interrupteurs de puissance bloqués : LDPC éteint

Maintenant, les interrupteurs de puissance sont bloqués. Le nombre d'interrupteurs de puissance, NSW, est 20% plus faible dans le circuit PD-SOI avec la solution Auto-DTMOS en comparaison avec le circuit PD-SOI avec la solution BC sans polarisation de body ou le circuit BULK.

Dans le cas d'un portage direct complet, du BULK vers le PD-SOI, et pour le même nombre NSW, la consommation statique en mode OFF est 10 fois plus élevée, pour la même tension d'alimentation, comme l'atteste la figure (issue des résultats mesurés sur les différents circuits de test de la librairie MTCMOS). Elle descend jusqu'à 5 fois lorsque la même en fréquence est ciblée ; rappelons que dans ce cas, la tension en BULK est de 1,4V et de 1,2V en PD-SOI. Le portage direct ne permet donc pas de mettre la technologie en avant en termes de consommation statique.



**figure : Mesure sur silicium de la résistance équivalente R<sub>off</sub> des 3 circuits de test d'interrupteurs de puissance FB, BC et Auto-DTMOS**

Par contre, si l'on compare les circuits conçus et fabriqués avec la plateforme PD-SOI intégrant des interrupteurs de puissance BC sans polarisation de body et Auto-DTMOS avec celui conçu et fabriqué en BULK, le circuit PD-SOI a une consommation statique du même ordre de grandeur que celui en BULK. Quand la solution BC sans polarisation de body présente un courant de fuite identique au BULK pour la même tension d'alimentation, la solution Auto-DTMOS permet de réduire de 20% celle-ci dans les mêmes conditions. Sachant que la solution Auto-DTMOS, à Ron équivalent, est implémentée avec 20% d'interrupteurs de puissance en moins.

Aux tensions d'alimentation qui amènent les mêmes fréquences, 360MHz sur les deux LDPC, BULK (1,4V) et PD-SOI (1,2V) ; lorsque les interrupteurs de puissance sont passants, la consommation statique est 2,8 fois supérieure pour la technologie PD-SOI en comparaison avec la technologie BULK. En désactivant les interrupteurs de puissance, la consommation statique est réduite par plus d'un facteur de 2 dans le pire cas (cf. tableau ), puisqu'elle est de 928nW pour le BULK et de 442nW pour le PD-SOI avec la solution Auto-DTMOS. La solution BC permet également de réduire la consommation statique, mais de manière moins agressive et en nécessitant plus de surface d'implémentation par rapport à la solution Auto-DTMOS.

<b>LDPC éteint</b>	<b>BULK 65LP</b>	<b>PD-SOI 65LP</b>	<b>PD-SOI vs BULK</b>
Static Power (nW)	928	442	<b>- 52.4 %</b>

**tableau : Comparaison à même fréquence maximale ciblée de la puissance statique du bloc LDPC éteint en BULK et PD-SOI avec la solution Auto-DTMOS**

Avec la solution « Auto-DTMOS », nous sommes donc capables de consommer moins pour les deux cas de figure, à même tension d'alimentation VDD (-20%) et même fréquence (-52,4%).

Cette réduction de la consommation statique du circuit PD-SOI, à même fréquence, a deux causes : un nombre d'interrupteurs de puissance unitaires réduit de 20% par rapport au circuit BULK et surtout une tension d'alimentation 14% plus faible. De plus, la solution Auto-DTMOS en libérant de l'espace permet d'assurer l'implémentation de plus de capacités de découplage, ce qui lisse au mieux le potentiel d'alimentation virtuelle VDDI et rend ainsi le bloc LDPC moins sensible aux appels en courant, Ion, et aux variations de la tension d'alimentation extérieure VDD.

#### 4.1. Conclusion

Les différents circuits démonstrateurs ont permis de mettre en avant le gain en vitesse et en consommation dynamique de la technologie PD-SOI par rapport à la technologie BULK en 65nm LP. Ils ont aussi permis de confirmer le portage direct de manière fonctionnelle avec un excellent rendement et de valider principalement la technique Auto-DTMOS en termes d'implémentation, de fonctionnement et de consommation statique. Cette technique permet de réduire la consommation statique, par plus de 2 (–52,4%), du domaine d'alimentation du LDPC lorsque celui est comparé à la technologie BULK et lorsque les mêmes fréquences sont ciblées. La plateforme commune 65nm LP BULK PD-SOI montre qu'il n'est pas nécessaire de développer entièrement un flot de conception PD-SOI. Il est préférable d'utiliser les outils BULK et de faire les validations nécessaires en PD-SOI pour la partie logique. La plateforme PD-SOI assure l'insertion finale des interrupteurs de puissance optimisés : Auto-DTMOS. Avec cette méthodologie et ces interrupteurs de puissance, la technologie PD-SOI devient une sérieuse alternative à la technologie BULK en termes de performance et de consommation électrique.

- [Ets'09] European Telecommunications Standard Institute (ETSI) "Digital Vidéo Broadcasting (DVB), EN 302 307 V1.2.1" [www.etsi.org](http://www.etsi.org), 2009
- [Gal'62] Gallager, R. "Low Density Parity Check Codes" *IRE Trans. Inform.Theory*, vol. IT-8, pp. 21, 1962
- [I3e'90] IEEE Standard for Information Technology "IEEE Std 1149.7" [www.standard.ieee.org](http://www.standard.ieee.org), 1990
- [I3e'09] IEEE Standard for Information Technology "IEEE Std 802.11n" [www.standard.ieee.org](http://www.standard.ieee.org), 2009
- [Mou'10] Moubdi, N.; Maurine, P.; Wilson, R.; Azemard, N.; Engels, S.; Rolindez, L. & Heinrich, V. "Voltage Scaling and Body Biasing Methodology for High Performance Hardwired LDPC" *IEEE International Conference on ICICDT*, P.82-85, 2010
- [Mut'99] Mutoh, S.; Shigematsu, S.; Gotoh, Y. & Konaka, S. "Design Method of MTCMOS Power Switch for Low-Voltage High-speed LSIs" *Design Automation Conference*, p.113-116, vol.1, 1999
- [Pot'09] Pottier, R. Tong, J. Hawkins, C. Kundu, R. Pelloie, J.-L. "ARM 1176 implementation in SOI 45nm technology and silicon measurement" *SOICONF*, p.1-4, 2009
- [Ura'05] Urard, P.; Yeo, E.; Paumier, L.; Georgelin, P.; Michel, T.; Lebars, V.; Lantreibecq, E. & Gupta, B. "A 135Mb/s DVB-S2 compliant Codec based on 64800b LDPC and BCH codes" *Digest of technical paper ISSCC*, p.446-447, 2005
- [Ura'08] Urard, P.; Paumier, L.; Heinrich, V.; Raina, N. & Chawla, N. "A 360mW 105Mb/s DVB-S2 Compliant Codec based on 64800b LDPC and BCH Codes enabling Satellite-Transmission Portable Devices" *Digest of technical paper ISSCC*, p.310-311, 2008



## 5. Les bascules de rétention en PD-SOI

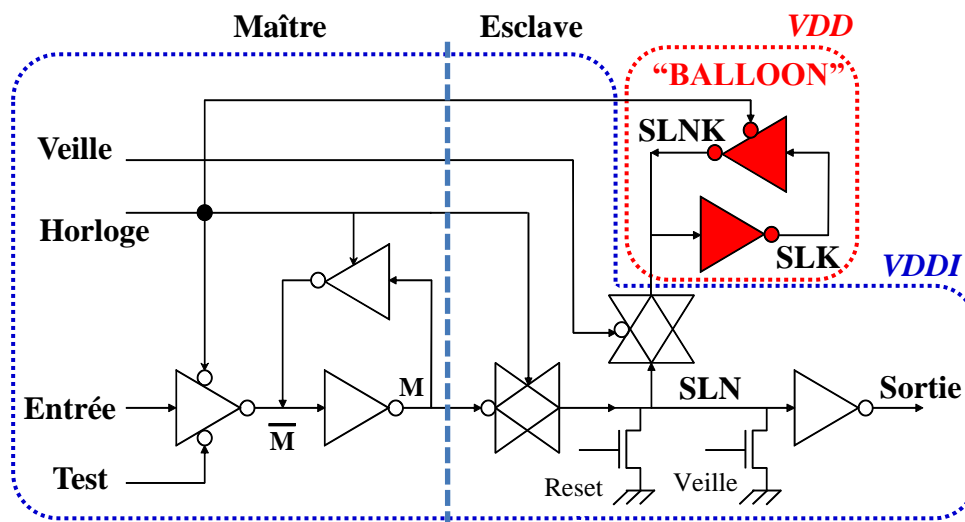
*Ce chapitre présente l'optimisation en PD-SOI du bloc devant être implémenté lorsque la technique MTCMOS avec interrupteurs de puissance est utilisée : la bascule de rétention. Ce bloc permet de conserver les états logiques lorsque les interrupteurs de puissance sont coupés et que l'alimentation virtuelle VDDI est à la masse. Il restitue le bon état logique après la phase de réveil, lorsque le circuit est à nouveau alimenté et opérationnel. La bascule de rétention a été optimisée dans un premier temps en 65nm BULK par STMicroelectronics, dans le but de supprimer un effet transitoire parasite et indésirable appelé « GLITCH ». Cet effet transitoire a un impact, dégradant la robustesse, plus important en PD-SOI qu'en BULK, d'où le fort intérêt de l'utiliser dans cette technologie. L'un des objectifs de cette thèse a été d'améliorer dans un second temps cette bascule de rétention sans effet transitoire en 65nm PD-SOI, en termes de consommation statique, de vitesse de transmission et de surface d'implémentation.*



## 5.1. Les Bascules de rétention

### 5.1.1. Description

Comme déjà indiqué dans le chapitre 2, l'utilisation d'interrupteurs de puissance pour couper les fuites au repos implique une perte de données (i.e. le contexte dans lequel se trouve le circuit logique avant l'extinction) si aucun mécanisme de sauvegarde n'est utilisé. C'est précisément le rôle des bascules de rétention (figure ). Celles-ci permettent de maintenir les états logiques lorsque le circuit est désalimenté [Hen'05] et de les restituer après la réalimentation.



**figure : Bascule de rétention**

Cette bascule de rétention est divisée en deux parties, la partie « maître » et la partie « esclave ». Ces deux parties se verrouillent de manière alternée. Le fonctionnement de cette bascule de rétention est classique, il assure l'échantillonnage du signal d'Entrée, au front montant du signal d'Horloge, et le maintien en Sortie durant tout le cycle d'Horloge.

La première partie, maître, devient passante lorsque le signal d'Horloge est au niveau bas, ce qui permet de propager la valeur du signal d'Entrée vers le nœud intermédiaire M. Elle se referme et se verrouille (« latch ») lorsque le signal d'Horloge passe au niveau haut. A l'opposé, la seconde partie, esclave, devient passante lorsque le signal d'Horloge est au niveau haut, ce qui permet de propager le signal du nœud intermédiaire M vers le nœud SLN et la Sortie.

La particularité de cette bascule de rétention est la sous-partie appelée « BALLOON » qui permet d'assurer la rétention. La sous-partie « BALLOON » est toujours alimentée, même quand les interrupteurs de puissance sont coupés, car elle est connectée à l'alimentation extérieure VDD, le reste de la bascule étant connecté à l'alimentation virtuelle VDDI. A chaque propagation du signal d'Entrée vers la Sortie de la bascule, la sous-partie « BALLOON » mémorise la valeur logique en transmission dans la partie esclave (celle du nœud SLN). Ce nœud se situe avant l'inverseur de sortie. Le « BALLOON » peut être représenté simplement par un point mémoire composé d'inverseurs montés tête-bêche.

Dans un mode de fonctionnement normal, le signal Test est à GND, les interrupteurs de puissance sont activés et le signal Veille, actif au niveau haut, a sa valeur à GND. Ce mode permet à la porte de passage (« pass-gate ») située entre les nœuds SLN et SLNK d'être passante et donc au « BALLOON » d'assurer un rafraichissement de sa mémoire périodiquement à chaque cycle d'horloge.

### 5.1.2. Rôle de la bascule de rétention

Son principal rôle est de conserver la donnée qui est située avant l'inverseur en sortie, lors de l'ouverture des interrupteurs de puissance et la mise à la masse de tout le domaine d'alimentation VDDI. Conserver cette donnée permet de la redistribuer soit à la sortie du circuit, pour les bascules situées à la fin, soit à la partie logique, pour les bascules placées au début. Il est important de constater que sans ce mode de rétention, deux phénomènes apparaîtraient après la remise sous tension.

Le premier est que les sorties de ces bascules auraient des valeurs aléatoires créées par l'instabilité que peut entraîner leur réveil. A titre d'exemple, il est impossible de prédire exactement, dans la figure , la valeur de la sortie du schéma électrique proposé si celui-ci passe d'un état éteint à un état alimenté et que son entrée est connectée à la sortie d'une bascule n'ayant pas de bloc de rétention.

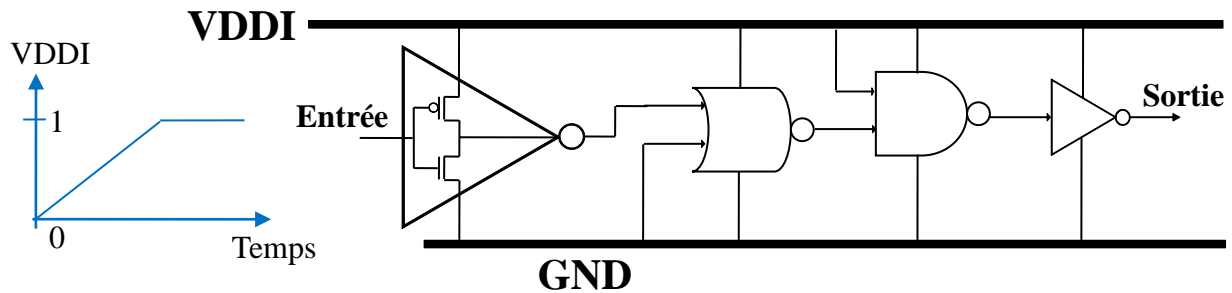


figure : Réveil d'un chemin logique

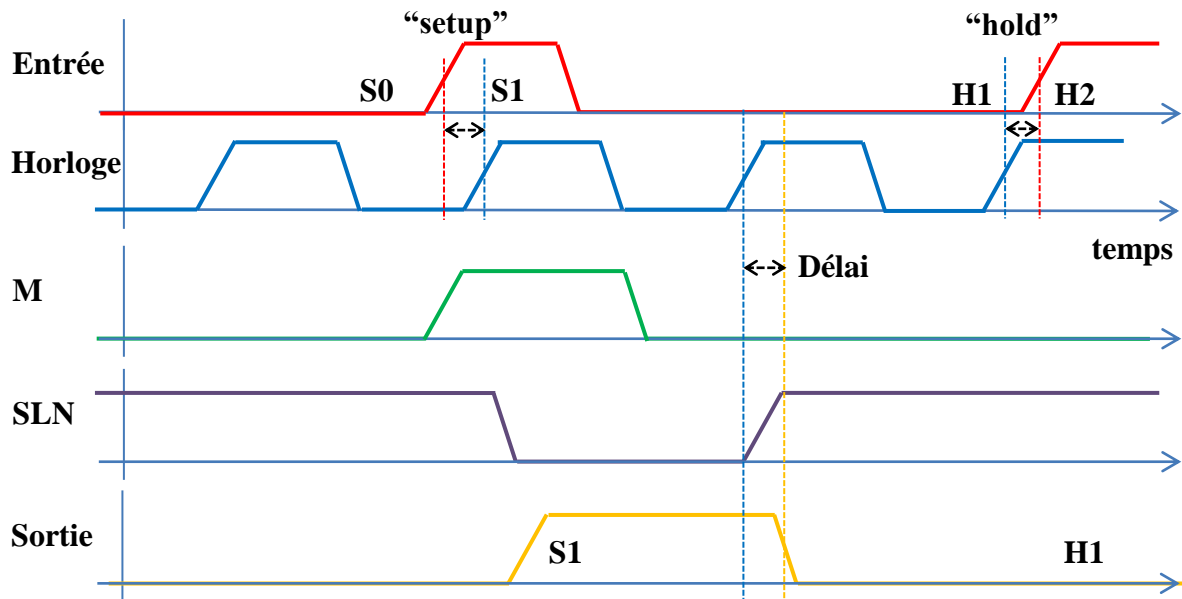
L'inconvénient serait alors de devoir faire une réinitialisation complète du circuit après chaque réveil. Cette phase de réinitialisation complexifie le circuit et la programmation de celui-ci mais plus encore introduit une perte de temps.

Le deuxième phénomène est lié au premier : le passage du mode ON à OFF des interrupteurs de puissance implique une interruption. Si, au cours ou même à la fin d'un calcul, le circuit se met dans un mode OFF, la perte des données entraînerait, une fois le circuit réveillé, de réexécuter l'opération une nouvelle fois depuis le début. L'interruption d'une tâche en cours de route est plutôt rare, mais avec des bascules de rétention, elle ne devient pas gênante puisque la tâche peut reprendre là où elle s'était arrêtée.

Avec la bascule de rétention, la mémorisation de la donnée n'empêche pas l'interruption d'un calcul mais permet de le poursuivre ou de le reprendre une fois les interrupteurs de puissance réactivés. Cela évite au calcul préalablement interrompu de passer dans un mode de réexécution. Ces deux phases (réinitialisation et réexécution) amènent de la complexité et requièrent un nombre important de cycle d'horloge. Ces bascules ont, en plus d'assurer un gain en termes de complexité et de mémorisation des données, un rôle clé dans le gain de temps.

### 5.1.3. Fonctionnement de la bascule de rétention

Mise à part la sous-partie « BALLOON », le comportement de cette bascule est semblable aux autres bascules, dites classiques, fonctionnant avec une seule tension d'alimentation. Quand la bascule de rétention est en mode normal, l'alimentation des parties maître et esclave ainsi que de la sous-partie « BALLOON » est la même.



**figure : Fonctionnement normal de la bascule de rétention**

La figure décrit le mode de fonctionnement normal équivalent à un mode sans technique « MTCMOS » avec interrupteurs de puissance (Veille=0). Il est normal de définir pour cette bascule de rétention son temps de propagation (Délai), son temps de maintien (« Hold ») et son temps de pré-positionnement (« Setup »). Le temps de maintien est le temps après le front montant de l'horloge durant lequel le signal d'entrée ne doit pas changer, au risque de prendre la valeur suivante. Le temps de pré-positionnement est le temps avant le coup d'horloge où le signal d'entrée doit être prêt et stable au risque de prendre la valeur précédente. Ces temps sont nécessaires à la conception des circuits logiques, ils permettent de vérifier que ceux-ci fonctionnent correctement.

Dans l'autre mode de fonctionnement (Veille=1), la porte de passage entre les nœuds SLN et SLNK devient bloquée, isolant le « BALLOON » du reste de la bascule (figure ).

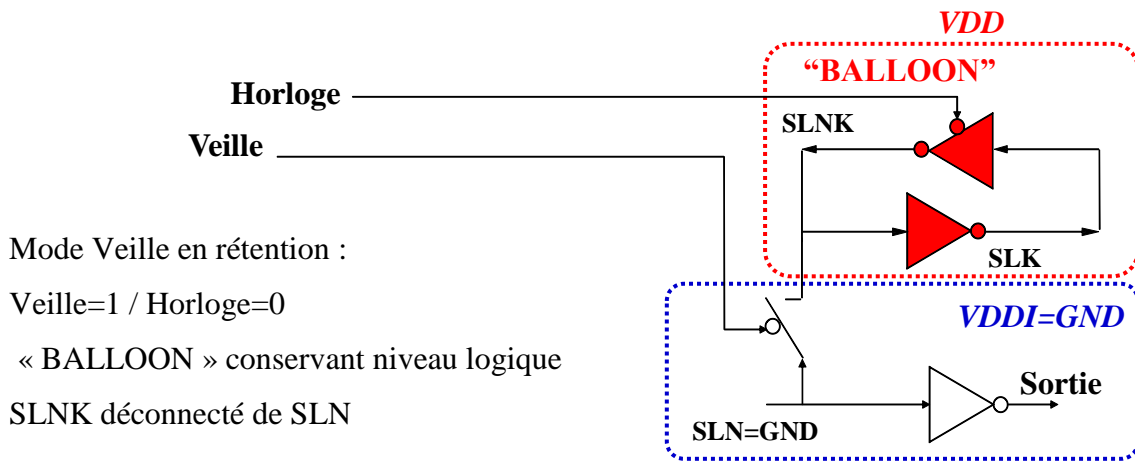


figure : Bascule de rétention en mode veille

#### 5.1.4. Environnement de la bascule de rétention

La bibliothèque MTCMOS contient les interrupteurs de puissance, les capacités de découplage et un contrôleur. Le contrôleur de cette bibliothèque (figure ) gère l'alimentation virtuelle VDDI et la commande des interrupteurs de puissance via SWVDD, mais également le signal de Veille et le signal d'Horloge. Par conséquent, le contrôleur pilote aussi le fonctionnement des bascules de rétention via ces deux derniers signaux.

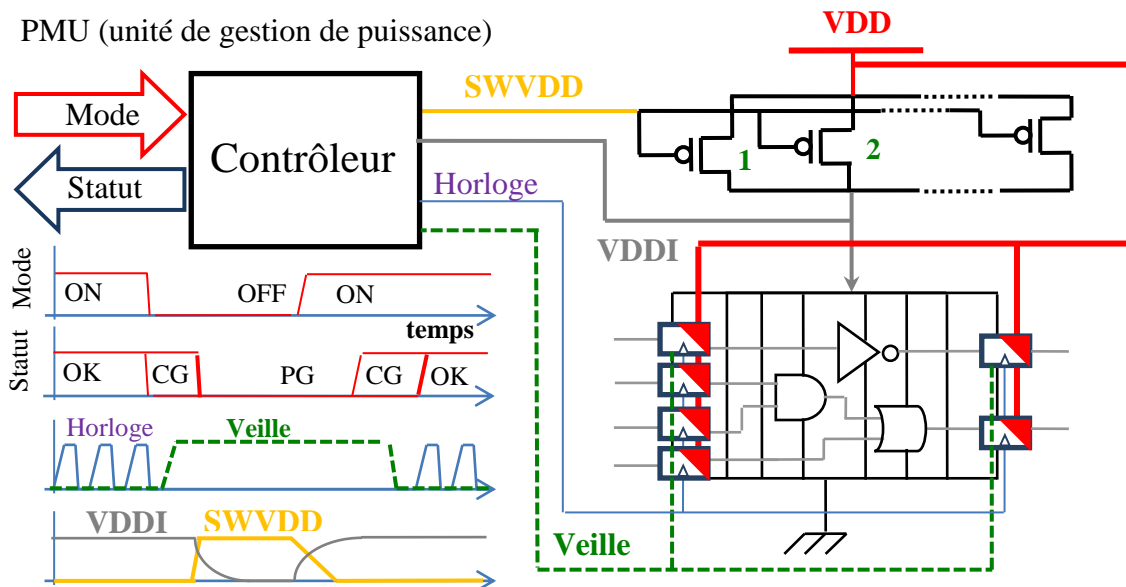


figure : technique MTCMOS avec bascules de rétention : schéma électrique et fonctionnement temporel

Les deux modes du contrôleur sont gérés par l'unité de gestion de puissance (PMU) : ON pour alimenter la partie logique, OFF pour la couper. En retour le contrôleur fournit des signaux de statut au PMU. Ces signaux reflètent les états dans lesquels sont la partie logique, l'arbre d'horloge et les interrupteurs de puissance. Le signal de statut OK signifie que la partie logique est alimentée et opérationnelle avec un arbre d'horloge activé et des interrupteurs de puissance passants. Le signal CG (« clock gating ») précise que l'arbre d'horloge est désactivé et que la partie logique n'est plus opérationnelle, même si elle est toujours alimentée par les interrupteurs de puissance. Le signal PG (« power gating ») est présent lorsque les interrupteurs de puissance sont coupés ; dans ce cas la partie logique est désalimentée et l'arbre d'horloge toujours désactivé. Le contrôleur est toujours dans un seul de ces modes.

Lors du passage du mode ON à OFF, le contrôleur coupe le signal d'Horloge (CG) de la partie logique ce qui permet d'activer le signal de Veille puis de couper les interrupteurs de puissance (PG). L'alimentation virtuelle VDDI coupée, la partie de rétention « BALLOON » reste la seule alimentée. L'isolation est faite en bloquant la porte de passage entre la sous-partie « BALLOON » et le reste de la bascule de rétention. L'isolation permet donc de mémoriser le dernier état logique du nœud SLN avant la coupure de l'alimentation VDDI par les interrupteurs de puissance. Lors du réveil, du mode OFF à ON, le contrôleur rallume progressivement les interrupteurs de puissance et participe au réveil en délivrant un courant à la partie logique. Il s'assure que l'alimentation virtuelle VDDI atteint une valeur proche de VDD, avant de libérer l'information retenue dans la partie de rétention « BALLOON ». Pour cela, le contrôleur réactive la porte de passage entre les nœuds SLN et SLNK, en mettant le signal Veille à un niveau bas. Quand les informations des différentes bascules de rétention sont toutes restituées, le contrôleur envoie un signal de statut au PMU pour lui dire que la partie logique est alimentée (CG), puis il active à nouveau le signal d'Horloge : le bloc logique devient opérationnel (OK).

### 5.1.5. La phase de réveil et l'effet transitoire parasite

Le point de rétention « BALLOON » peut mémoriser aussi bien le niveau logique 0 que le niveau logique 1. Dans le cas où le point de rétention mémorise 0, et lorsque le signal Veille réactive la porte de passage, la restitution n'engendre aucun conflit puisque SLNK et SLN sont au même potentiel, GND. Par contre, dans le cas où le point en rétention est au niveau logique 1 (VDD), la restitution induit un conflit puisque SLNK qui est à VDD et SLN

qui est à GND se voient tous deux connectés l'un à l'autre via la porte de passage. Le but est que SLNK restitue à SLN le niveau logique 1 qu'il avait mémorisé, créant un transfert de charge entre ces deux nœuds. Ce transfert de charges entraîne un effet transitoire parasite indésirable appelé « glitch » [Akl'08]. Le nœud SLNK, en délivrant au nœud SLN son niveau logique 1 mémorisé, voit transitoirement son potentiel chuter.

Dans la majorité des conditions de fonctionnement (pour toute la gamme de tension d'alimentation, de température et de qualité de procédé), cet effet parasite transitoire est sans conséquence sur la restitution du niveau logique 1 du nœud SLNK au nœud SLN. Mais il peut arriver, dans des conditions exceptionnelles, que la chute du potentiel SLNK fasse basculer le signal de sortie de l'inverseur auquel il est connecté, SLK. Dans ce cas, l'effet transitoire parasite n'est pas sans conséquence puisqu'il induit une faute en inversant SLK, de GND à VDD, ce qui force dans un second temps à inverser SLNK, de VDD à GND. Cela revient à rendre la rétention inutile, puisqu'elle perd son information. En 65nm PD-SOI, cette faute est plus souvent observable qu'en BULK, en raison du pic de courant supérieur et du plus faible découplage capacitif naturel du MOS FB par rapport au MOS BULK.

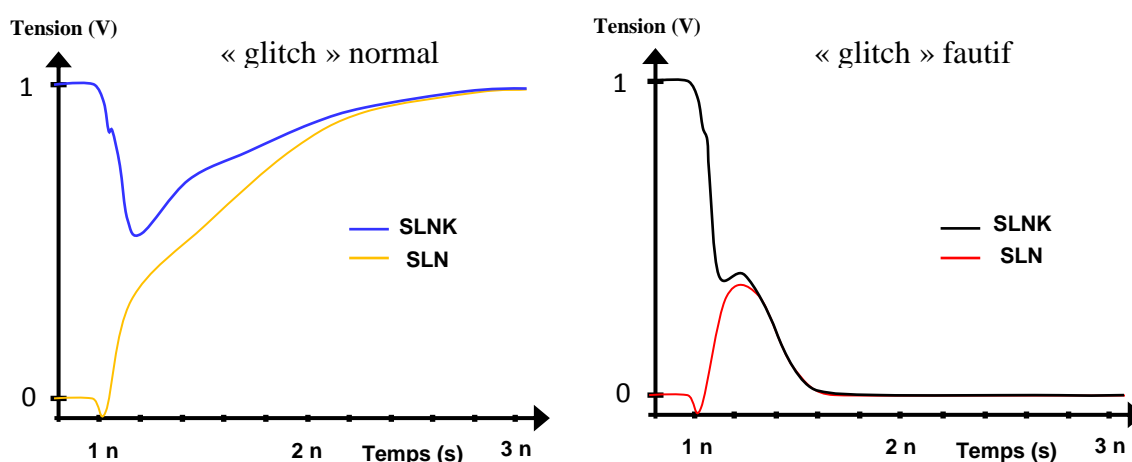


figure : « glitch » normal et entraînant une faute à 125°C lors de la restitution

Dans la figure , les deux cas de figure de réveil sont représentés. Le premier, normal, montre un réveil où le conflit entre les nœuds SLNK et SLN n'induit pas d'inversion du potentiel SLNK mais juste un abaissement de son potentiel. Le nœud SLN prend finalement le niveau logique 1 stocké sur le nœud SLNK (VDD). Le second cas, par contre, montre que le conflit abaisse suffisamment le potentiel SLNK pour faire basculer le point mémoire de la sous-partie « BALLOON ». Dans ce cas, SLNK et SLN sont tirés vers la masse. Ce phénomène apparaît dans des conditions extrêmement défavorables et peu probables. Mais

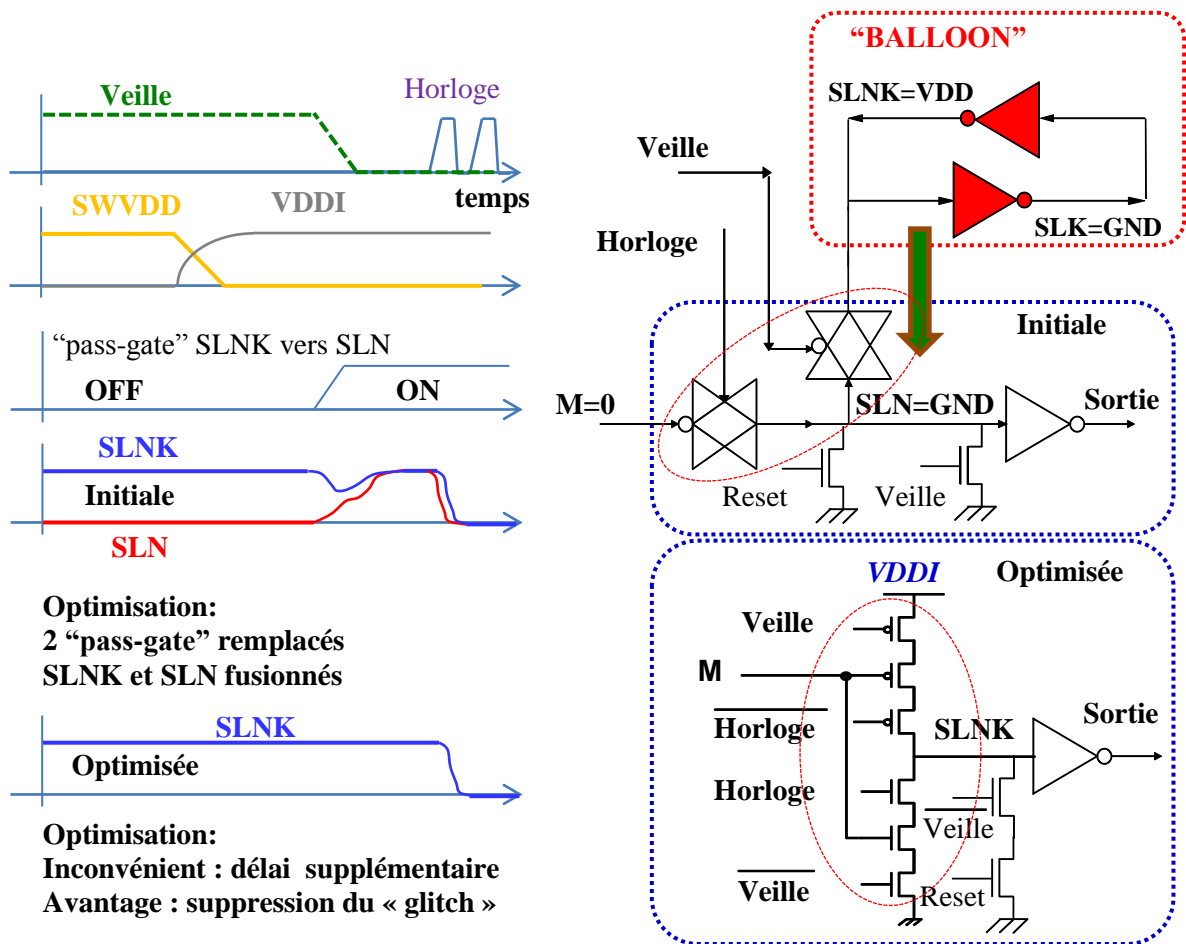
que ce soit en technologie BULK ou en PD-SOI, cet effet transitoire parasite peut entraîner une mauvaise relecture de la mémoire du « BALLOON », ce qui ne rend pas la bascule de rétention assez robuste. D'où le fort intérêt de travailler avec une bascule de rétention sans effet transitoire parasite en 65nm PD-SOI.

A STMicroelectronics, des concepteurs se sont donc attachés à supprimer ce phénomène parasite dans une bascule de rétention optimisée en technologie BULK. Dans le cadre de mes travaux, j'ai dérivé cette bascule de rétention optimisée en technologie PD-SOI et me suis focalisé sur l'amélioration de celle-ci en termes de consommation statique, de performances et de surface. Au niveau de la conception, des transistors à différentes tensions de seuil sont utilisés : (1) de type LVt ou SVt, pour leur vitesse, dans les parties alimentées par les interrupteurs de puissance ; et (2) de type HVt, pour leur faible courant de fuite, dans la sous-partie toujours alimentée « BALLOON ». En PD-SOI, sont également utilisés des transistors de type FB ou BC.

### 5.2. Suppression de l'effet transitoire

L'effet transitoire parasite est dû au conflit lors du réveil entre le potentiel SLNK qui a mémorisé le niveau logique 1 (VDD) et le nœud SLN qui est fixé à la masse durant cette période. Il est la conséquence du passage de l'état éteint à l'état passant de la porte de transfert, ce qui permet de libérer les charges de SLNK vers SLN.





**figure : Phase de réveil avec solution initiale et optimisée**

L'optimisation de la bascule de rétention consiste à remplacer les portes de transfert par une cellule inverseuse (figure ) contrôlée par les signaux de Veille et d'Horloge, qui permet d'isoler le nœud SLNK ou de verrouiller le signal de la partie maître : **M**.

Avec la solution initiale, lorsque la bascule de rétention était dans un mode de veille, il était conseillé de fixer le potentiel SLN à la masse, afin d'éviter des niveaux intermédiaires sur ce nœud pouvant entraîner un courant de court-circuit lors de la phase de réveil. De plus, c'était ce nœud qu'il fallait remettre à zéro lors de réinitialisation. Il fallait donc deux MOS en parallèle sur ce nœud, l'un commandé par le signal Veille, l'autre par le signal Reset. Avec la solution optimisée, les nœuds SLNK et SLN sont fusionnés. SLNK devient à la fois le nœud de rétention et le nœud de propagation du signal du nœud M vers la Sortie.

Le nœud SLN n'existant plus et le nœud SLNK ne pouvant être remis à zéro qu'en mode normal, les deux MOS commandés par Reset et Veille, en parallèle dans la solution initiale, sont remplacés par deux MOS empilés avec la solution optimisée. L'un est contrôlé par le signal Reset et l'autre par l'inverse du signal Veille. En effet, la réinitialisation se fait lorsque la bascule de rétention n'est pas en veille. Le fait que ces deux MOS en parallèle

deviennent deux transistors empilés amène aussi une réduction du courant de fuite. Cette optimisation introduit néanmoins un délai supplémentaire mais assure un fonctionnement sans effet transitoire parasite et donc beaucoup plus robuste.

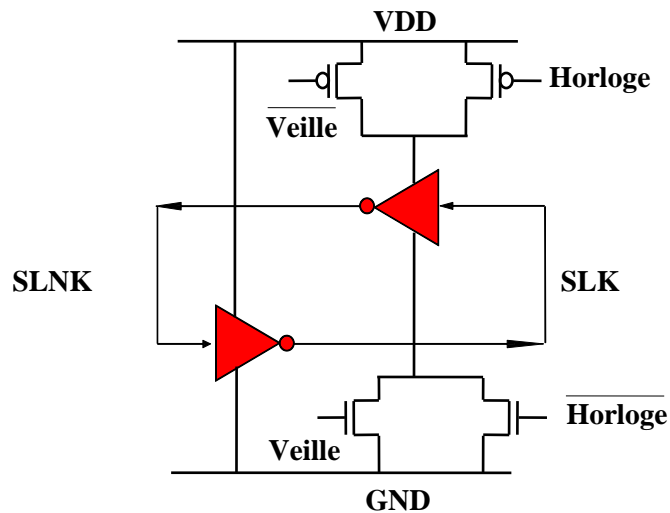
Avant d'analyser l'impact sur le délai, l'optimisation va se poursuivre en PD-SOI en proposant plusieurs solutions dans le but non seulement de supprimer le phénomène parasite mais également de réduire au maximum le délai interne, la consommation statique et la surface d'implémentation.

### 5.3. Amélioration pour la technologie PD-SOI

#### 5.3.1. Etude de structures pour réduire le courant de fuite

En PD-SOI, les transistors FB utilisés dans la bascule de rétention sont plus fuyants qu'en BULK à cause de l'effet flottant du body qui réduit la tension de seuil des transistors et augmente le courant sous le seuil. L'optimisation en technologie PD-SOI de la consommation en rétention consiste à adapter le schéma en introduisant des transistors BC ou en empilant des MOS de type FB. Les transistors FB des parties alimentées à VDDI sont mis à la masse lorsque les interrupteurs de puissance sont éteints. De plus, en termes de performances, ils sont meilleurs que les transistors BC ou FB empilés ; ils n'ont donc pas besoin d'être remplacés. La réduction de la consommation statique doit se faire sur la sous-partie toujours alimentée, « BALLOON ».

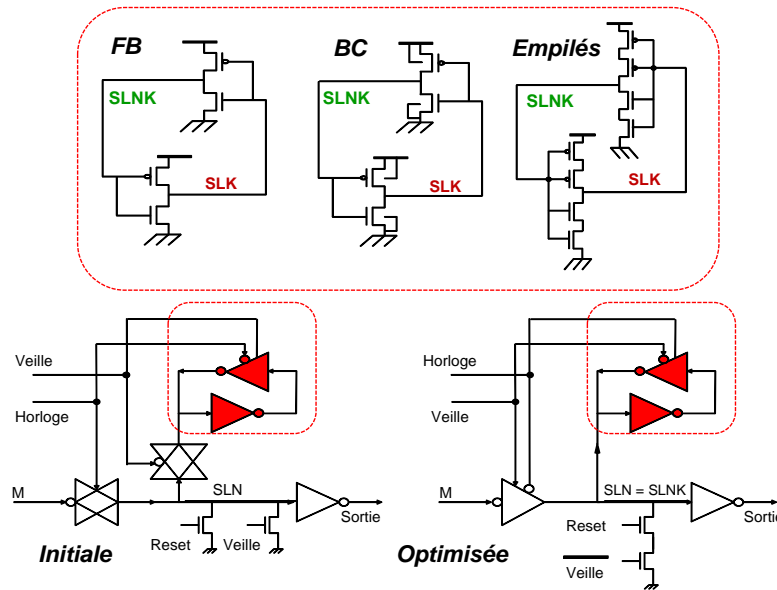
En BULK ou en PD-SOI, la sous-partie « BALLOON » est de type HVt alors que le reste de la bascule de rétention est de type SVt et LVt. Dans l'optimisation, seule la partie mémoire du « BALLOON » est optimisée, à savoir les deux inverseurs montés tête-bêche. L'inverseur ayant son entrée connectée à SLK est commandé par le signal d'Horloge et le signal de Veille. En mode de rétention, le signal d'Horloge est à 0 et le signal Veille à 1 (figure ).



**figure : Activation de la sous-partie « BALLOON »**

Les signaux Veille et Horloge sont connectés à des MOS de type P et N et ont pour rôle d'alimenter ou de laisser le nœud SLNK flottant. En mode de veille ou lorsque le signal d'Horloge est bas, ces transistors sont passants et ne permettent pas de réduire le courant de fuite. Leur activation permet un rafraichissement de la donnée sur SLNK, une fois la partie esclave isolée de la partie maître.

Deux optimisations en consommation sont proposées pour les inverseurs montés tête-bêche de cette sous-partie, en se basant sur l'utilisation de transistors BC ou l'utilisation d'un empilement de deux MOS FB (figure ), et en considérant le cas où l'inverseur du haut est activé (Veille=1 & Horloge= 0).



**figure : Optimisation de la sous-partie « BALLOON » en consommation statique**

Nous avons comparé plusieurs schémas alternatifs, en considérant leurs consommations statiques en veille, leurs surfaces et leurs comportements dynamiques. La solution initiale est celle qui possède toujours un effet transitoire, la solution optimisée est celle sans ce phénomène. Les solutions initiale et optimisée sont déclinées avec des « BALLOONS » conçus avec des transistors de type FB, BC et FB empilés. Nous cherchons ainsi à réduire les fuites en utilisant des transistors à Body Contacté, ou en exploitant les propriétés de l'empilement de transistors (voir chapitre 2). Le nom de chaque solution est composé de deux mots, le premier définit si la solution est initiale ou optimisée, le second nous donne le type de « BALLOON », FB, BC ou FB empilés. La solution initiale avec un « BALLOON » qui n'a pas été modifié est de type FB (Initiale\_FB), elle fera office de référence.

### 5.3.2. Comparaison des différentes implémentations

#### 5.3.2.1. Surface

La solution optimisée permettant de supprimer l'effet parasite sans modification des « BALLOONS » (Optimisée\_FB) ne requiert pas plus de surface que la solution de référence. Par contre, le passage de la solution Optimisée\_FB vers un « BALLOON » de type BC (Optimisée\_BC) ou FB empilés (Optimisée\_Empilés) introduit une surface supplémentaire (tableau ).

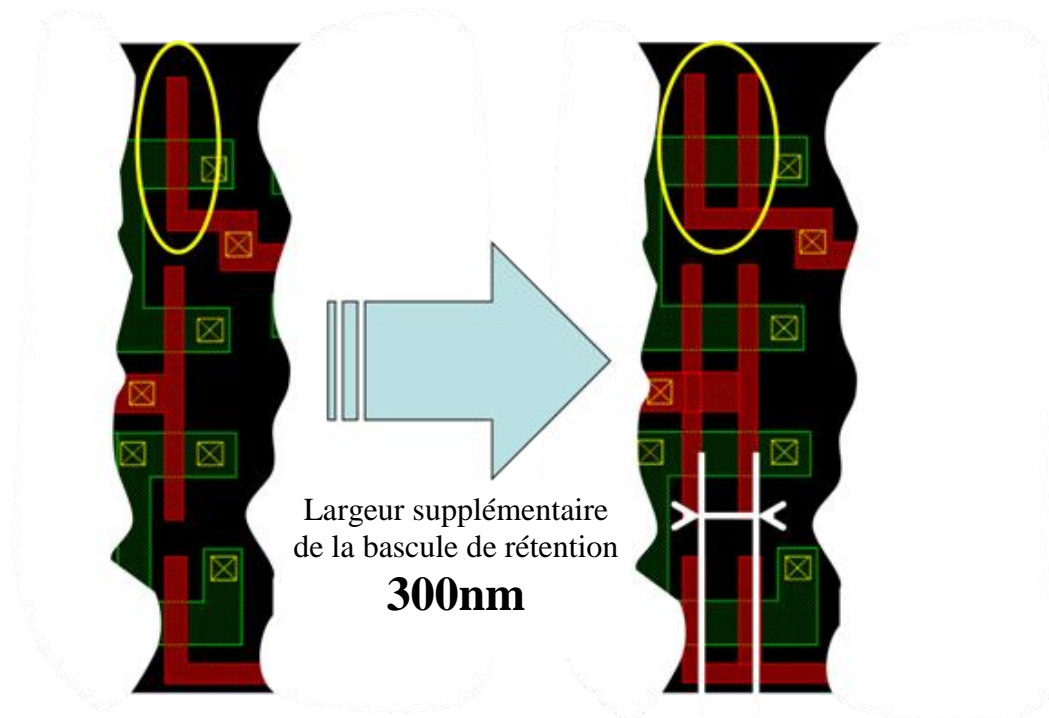
Solution	Initiale FB "BALLOON"	Optimisée FB "BALLOON"	Initiale BC "BALLOON"	Optimisée BC "BALLOON"	Optimisée Empilés "BALLOON"
Dégradation de surface	Référence	<b>0 %</b>	<b>+ 25 %</b>	<b>+ 25 %</b>	<b>+ 6 %</b>

**tableau : Comparaison de la surface des bascules de rétention**

Les solutions utilisant un « BALLOON » de type BC nécessitent une surface de silicium 25% plus grande que la solution de référence. La solution Optimisée\_Empilés a une surface seulement 6% plus élevée.

Le transistor BC induit une surface bien supérieure au transistor FB pour des largeurs de grille minimale, car la prise « body » est fixe et indépendante de la largeur du MOS. Ainsi, l'impact d'une prise « body » utilisée sur les transistors de petite taille de la bascule de rétention est relativement important ( $\approx 50\%$ ). Leur placement contraint d'espacer la prise « body » des autres MOS car le dopage de celui-ci est en opposition avec les drains et sources des autres MOS. Il convient en plus de connecter les prises « body ». L'implémentation de transistors BC induit aussi de router la cellule avec un niveau de métal 2 supplémentaire alors qu'avec la solution initiale, le métal 1 était suffisant pour router la bascule de rétention.

A l'inverse, dans le cas de l'empilement de transistors, il n'y a pas de contrainte d'implémentation ni de routage. Il faut juste espacer la zone entre le drain et la source des MOS des inverseurs en question, pour y insérer une deuxième grille, comme le montre la figure . En termes de surface, mieux vaut donc utiliser un empilement de transistors dans les blocs logiques où sont intégrés des MOS à forte densité, plutôt qu'utiliser des transistors BC.



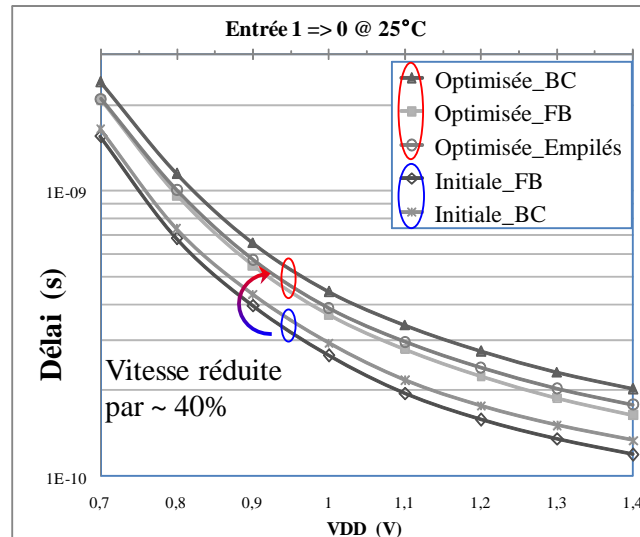
**figure : Passage des 2 inverseurs du « BALLOON » de FB à FB empilés**

Avec l'empilement de transistors une partie de la de la cellule est décalée de 300nm, ce qui induit donc une longueur totale environ 6% plus élevée. Cela n'induit aucun changement dans le routage interne de la bascule de rétention. Avec des transistors de type BC, il faut tout revoir, le placement des MOS ainsi que leur routage qui implique le routage des « body ». Les transistors BC sont véritablement désavantageux en termes de surface et contraignants à implémenter dans les blocs des bibliothèques numériques.

#### 5.3.2.2. Vitesse

Cinq solutions sont comparées en termes de vitesse, Initiale\_FB (référence), Initiale\_BC, Optimisée\_FB, Optimisée\_BC et Optimisée\_Empilés. Les deux premières solutions sont conçues avec les portes de passage et l'effet transitoire parasite et avec un « BALLOON » de type FB et BC. Les trois autres sont basées sur la solution optimisée pour supprimer le phénomène transitoire (« glitch ») avec un « BALLOON » de type FB, BC et FB empilés.

Le passage de la solution initiale à la solution optimisée induit un délai supplémentaire dû au remplacement des portes de passage par l'inverseur contrôlé par les signaux Veille et Horloge. Le « BALLOON » de type FB apporte une capacité inférieure sur SLNK, par rapport aux deux autres solutions, impactant le délai interne de la bascule de rétention.



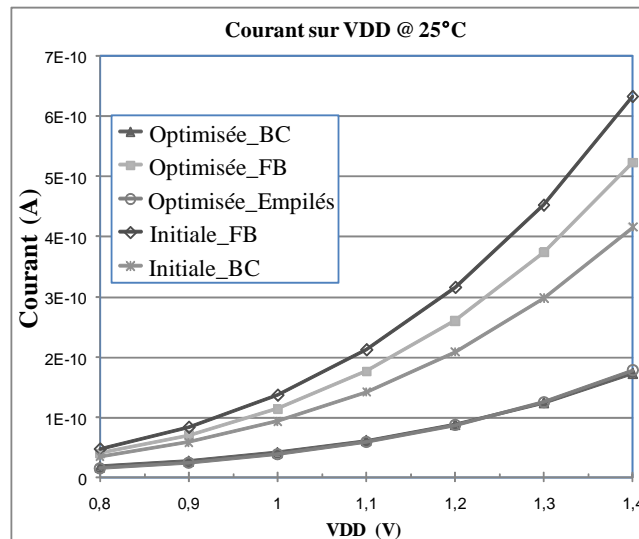
**figure : Comparaison du délai de la bascule de rétention selon les 5 cas proposés**

La figure montre la dégradation de vitesse qu'introduit l'élimination du transitoire parasite de la bascule de rétention. Ce délai supplémentaire provient du remplacement des portes de passage par un inverseur à trois MOS empilés (i.e. un MOS pour le signal Horloge, un MOS pour le signal de Veille et un autre pour le signal M). Parmi les implémentations supprimant l'effet parasite, la moins critique en termes de vitesse est celle à base de transistors FB, suivie de celle utilisant des MOS FB empilés. La plus lente est l'implémentation à base de transistors BC. Cette dernière amène une capacité équivalente sur le nœud SLNK plus importante que l'implémentation avec transistors FB empilés, car un MOS BC aura une largeur minimale de 300nm alors que le MOS FB peut, lui, descendre jusqu'à 120nm. De plus, la prise « body » obtenue par une astuce « layout » requiert plus de surface de grille, augmentant la capacité vue sur le nœud SLNK. La mise en série des deux MOS empilés sans contact augmente la capacité de grille d'environ 50% alors que le BC l'augmente d'un facteur supérieur à 3. La solution Optimisée\_FB est la plus rapide mais reste celle possédant le plus de courant de fuite. Pour de faibles courants de fuite, l'implémentation Optimisée\_Empilés est la plus avantageuse en termes de vitesse et de surface d'implémentation.

## 5.3.2.3. Consommation statique

Le courant de fuite se mesure lorsque les interrupteurs de puissance sont en veille. Les solutions optimisées montrent une première réduction du courant de fuite par rapport aux solutions initiales et pour le même type de « BALLOON ». En effet, la solution Optimisée\_FB consomme moins que la solution Initiale\_FB et il en est de même pour la solution Optimisée\_BC vis-à-vis de la solution Initiale\_BC. Cela est dû à l'empilement de transistors dans l'inverseur contrôlé par les signaux de Veille et d'Horloge, mais également à la mise en série des deux MOS de gestion du potentiel SLNK contrôlés par le signal Reset et par l'inverse du signal Veille (figure ).

Les solutions Optimisée\_BC et Optimisée\_Empilés sont celles qui consomment le moins, de manière quasi identique à 25°C et pour une tension d'alimentation VDD allant de 0,8V à 1,4V (figure ). Cette réduction des courants de fuite s'explique pour les transistors FB empilés pour une différence de potentiel divisée par 2 et un effet de body flottant fortement réduit ; de plus le MOS du haut de l'empilement a un V<sub>gs</sub> négatif. Le transistor BC, a lui son body polarisé à la source et possède donc une tension de seuil plus élevée.

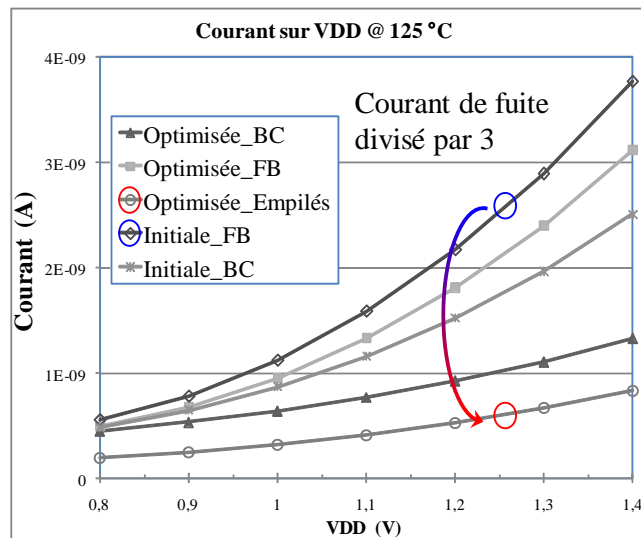


**figure : Comparaison de la consommation statique en veille selon les 5 cas proposés à 25°C**

Par contre, à plus haute température, la solution Optimisée\_Empilés se démarque de manière significative de la solution Optimisée\_BC (figure ). Avec l'augmentation de la température, le courant sous le seuil du transistor BC croît plus rapidement que celui des deux MOS empilés. Il est rare d'avoir cette température dans un mode de veille, néanmoins l'étude



à 125°C permet de déterminer la solution la plus intéressante, puisque les deux implémentations considérées présentent les mêmes courants de fuite à 25°C.



**figure : Comparaison de la consommation statique en veille selon les 5 cas proposés à 125°C**

#### 5.4. Implémentation optimale

Grâce aux comparaisons précédentes, il est possible de déterminer que l'implémentation de la sous-partie « BALLOON » à base de transistors FB empilés est celle présentant le meilleur compromis de surface, de courant de fuite et de vitesse. C'est donc celle-ci qui est retenue pour les futures implémentations de circuits incluant des interrupteurs de puissance et nécessitant des bascules de rétention optimisées. Cette implémentation peut également être exploitée en technologie BULK, avec le même bénéfice de suppression du transitoire parasite. Cela permet notamment de concevoir des blocs logiques numériques entièrement en BULK avec la bascule de rétention proposée et de réaliser un portage direct en technologie PD-SOI, en remplaçant uniquement les interrupteurs de puissance et autres blocs spécifiques à la technologie PD-SOI comme les Entrées/Sorties.

- [Akl'08] Akl, C.J. & Bayoumi, M.A. "Reducing Wakeup Latency and Energy of MTCMOS Circuits via Keeper Insertion" ISLPED'08, 2008
- [Hen'05] Henzler, S.; Nirschi, T.; Pacha, C.; Spindler, P. et al. "Dynamic state-retention flip-flop for fine-grained sleep-transistor scheme" European Solid-State Circuits Conference, p.145-148, 2005



## ➤ Conclusion

Les travaux de cette thèse ont cherché à démontrer tous les avantages fournis par la technologie 65nm LP PD-SOI en termes de vitesse et de consommation électrique. Une solution spécifique a notamment été proposée pour la réduction de la consommation statique. Les différents circuits démonstrateurs présentés dans ce mémoire sont basés sur une approche industrielle et ont été conçus en suivant le même flot de conception et de fabrication. Ces circuits ont montré tous les bénéfices apportés par la technologie PD-SOI par rapport à la technologie BULK. L'association de la haute performance des transistors à body flottant (FB) et de la basse consommation des interrupteurs de puissance de type body contacté (BC) en font une technologie mature idéale pour de nombreuses applications. Nous avons montré qu'à même tension, cette technologie apporte un gain d'environ 20% en fréquence typique, par rapport à son homologue BULK. Dans un autre mode opératoire, à plus faible tension d'alimentation et à une fréquence identique au BULK, la technologie PD-SOI permet de réduire de 30% la consommation dynamique. Quand le premier cas est favorable aux applications fixes, le second cas l'est pour les applications mobiles. Ces mesures ont été réalisées sur des circuits ayant une partie logique complexe et représentative de ce que propose le marché du semi-conducteur pour les applications fixes et mobiles. Afin de réduire au maximum la consommation électrique en veille, une technique de réduction de la consommation statique a été proposée. Cette technique entraîne, en PD-SOI et dans un mode veille, une consommation de fuite divisée par deux par rapport à la technologie BULK lorsque les mêmes fréquences sont ciblées.

La connaissance de la technologie SOI et des différentes techniques de réduction de la consommation électrique, ainsi que des contraintes de plus en plus élevées en termes de surface, m'ont permis d'évaluer la technique MTCMOS et de proposer un nouveau facteur de sélection des interrupteurs de puissance. Ce critère de sélection permet de mettre en avant la solution ayant le meilleur compromis en termes de consommation statique, de performances et de surface. La solution considérée comme la solution optimale utilise un interrupteur de puissance de type BC avec une longueur de grille de 200nm. A partir de là, une solution innovante appelée Auto-DTMOS, basée sur l'optimisation de cette solution optimale, a été conçue et fabriquée. Cette optimisation est réalisée grâce à une polarisation auto-adaptative du body, réduisant la résistance équivalente en mode passant des interrupteurs de puissance d'au

moins 20%. Dans un premier temps, cette solution a été implémentée, de manière isolée, sur des circuits de test et comparée à la solution optimale et la solution de technologie BULK. Ainsi validée, la solution a ensuite été implémentée et associée à une partie logique complexe et fortement fuyante. Les différents circuits réalisés montrent un gain technologique du PD-SOI par rapport au BULK. Sans technique de réduction de la consommation statique, le courant de fuite de la technologie PD-SOI est 3 à 10 fois supérieur à celui de la technologie BULK. Avec des interrupteurs de puissance, la réduction de la consommation statique est réduite de 2 à 3 décades, et devient même inférieure à celle du circuit BULK.

En 65nm PD-SOI, la solution proposée Auto-DTMOS permet de réduire de 20% la surface et la consommation statique, à même tension d'alimentation et pour la même résistance équivalente du réseau d'interrupteurs de puissance. A une tension d'alimentation plus faible, lorsque les mêmes fréquences sont ciblées en PD-SOI et en BULK, la consommation statique des interrupteurs de puissance est plus de deux fois plus faible. Cette solution permet aussi d'insérer plus de capacités de découplage, ce qui lisse la tension d'alimentation.

La bascule de rétention, indispensable à associer à la technique MTCMOS a été optimisée en PD-SOI. Elle offre l'avantage de n'être implémentée qu'avec des transistors de type FB et de ne pas présenter d'effet transitoire parasite. Par rapport à son équivalent en BULK, elle a un courant de fuite en veille identique pour seulement 6% de surface supplémentaire.

Parmi les perspectives, reste en particulier l'optimisation du contrôleur qui permet de commander la mise en marche/arrêt des interrupteurs de puissance. Le contrôleur, utilisé dans tous nos circuits, est issu directement de la bibliothèque BULK. Il est donc composé de MOS de type FB en PD-SOI. L'optimisation en consommation consisterait à remplacer les MOS de type FB par des transistors de type FB empilés ou BC pour amener le contrôleur à fuir au minimum. Il conviendrait également de vérifier et corriger la robustesse des blocs analogiques, en prenant en compte des phénomènes spécifiques au SOI comme l'effet « Kink » ou l'auto-échauffement.

Les récents progrès et l'évolution des technologies ont amené à passer d'une technologie SOI partiellement désertée (PD) à une technologie entièrement désertée (FD), plus efficace à partir du nœud 28nm. Cette technologie fonctionne avec des transistors sans dopage de canal, réduisant les problèmes de variation de procédé et donc la dispersion des

performances. Elle possède de meilleures caractéristiques électriques, notamment la pente sous le seuil, ce qui réduit intrinsèquement les courants de fuite des transistors. Néanmoins, l'intégration à très grande échelle impose toujours de couper les courants de fuite résiduels par des interrupteurs de puissance. Ainsi, les techniques de conception et la méthodologie développées au cours de cette thèse peuvent être appliquées à la technologie FD-SOI, nécessitant seulement quelques adaptations. Par exemple, la solution Auto-DTMOS pourrait être adaptée en exploitant la technologie FD-SOI dite UTBB (Ultra-Thin Body and Box), qui permet de moduler la tension de seuil des transistors par une polarisation face arrière.

La technologie SOI a longtemps été freinée par son coût et par les particularités de sa conception. La tranche de silicium sur isolant est en effet plus chère que la tranche de silicium et la conception d'un circuit en SOI impliquait auparavant une connaissance pointilleuse de cette technologie. Les contributions apportées par les travaux de cette thèse sont des éléments de bibliothèque dont la mise en œuvre est la même que leurs équivalents en CMOS BULK et ne nécessite pas de connaissance particulière du SOI.

En conclusion, les travaux de cette thèse ont conduit à la proposition d'un nouveau facteur de sélection des interrupteurs de puissance, à l'optimisation de la bascule de rétention en PD-SOI et principalement à l'invention de la solution Auto-DTMOS. Cette solution est générique aux technologies PD-SOI et semble aussi dérivable à la technologie FD-SOI. La technologie PD-SOI est plus avantageuse que la technologie BULK, en termes de vitesse et de consommation dynamique. Elle devient aussi, grâce aux travaux de cette thèse, une sérieuse alternative à la technologie BULK en termes de consommation statique.



## Brevet et publications

### *Brevet déposé*

« Circuit de polarisation dynamique du substrat d'un transistor » FR10/03656 US13/232.529

### *Conférences internationales IEEE avec acte*

- [Dag'09] D'agostino, C.; Le Coz, J.; Flatresse, P. Beigne, E. & Belleville, M. "An Accurate Approach for Statistical Estimation of Leakage Current Considering Multi-Parameter Process Variations in Nanometer CMOS Technologies", ESSDERC, p.427-430, 2009
- [Lec'10b] Le Coz, J.; Valentian, A.; Flatresse, P. & Belleville, M. "Power Switch Optimization and Sizing in 65nm PD-SOI Considering Supply Voltage Noise", ICICDT, p.186-189, 2010
- [Lec'11a] Le Coz, J.; Flatresse, P.; Engels, S. ; Valentian, A.; Belleville, M.; Raynaud, C.; Croain, D. & Urard, P. "Comparison of 65nm LP bulk and LP PD-SOI with adaptive power gate body bias for an LDPC codec", ISSCC, p. 336-337, 2011
- [Lec'11b] Le Coz, J.; Flatresse, P.; Clerc, S. ; Valentian, A. & Belleville, M. "65nm PD-SOI Glitch-Free Retention Flip-Flop for MTCMOS Power Switch applications", ICICDT, p.1-4, 2011

### *Conférence européenne avec acte*

- [Lec'10a] Le Coz, J.; Valentian, A.; Flatresse, P. & Belleville, M. "Power Switch Optimization in 65nm PDSOI Considering Physical Implementation Constraints", EuroSOI, 2010





# Glossaire

BC	Body contacté : Transistor à prise body
BOX	« buried oxide » oxyde enterré
CMOS	« complementary metal oxide semiconductor » transistors complémentaires
DIBL	« Drain Induced Barrier Lowering » Abaissement de la barrière de potentiel induit par le drain
DTMOS	« Dynamic Threshold MOS » Transistor à tension de seuil dynamique
FB	« Floating Body » Transistor à body flottant
FBB	« Forward Body Biasing » Polarisation en direct de body
FD-SOI	« Fully Depleted » Technologie Entièrement Désertée SOI
GIDL	« Gate Induced Drain Leakage » Courants de fuite de drain induits par la grille
GLITCH	effet transitoire parasite
GND	masse d'alimentation
Hold	temps de maintien
HP	« High Performance » Hautes performances
HV <sub>t</sub>	Haute tension de seuil
IDDQ	courant de fuite de la partie logique alimentée
I/O	« Input / Output » Bloc d'Entrée / Sortie
I <sub>off</sub>	courant de fuite
I <sub>on</sub>	courant dynamique
IP	« Intellectual property » bloc avec propriété intellectuelle
Latch	verrou
LDPC	« Low Density Parity Check » vérificateur de parité à faible densité
LP	« Low-Power » Faible puissance
LV <sub>t</sub>	Basse tension de seuil
MOS	« metal oxide semiconductor » transistor
MTCMOS	« Multi Threshold CMOS » CMOS à tension de seuil multiple
Multi-VT	implémentation à l'aide de transistors avec plusieurs tensions de seuil
NSW	« Number of switches » Nombre d'interrupteurs de puissance
PD-SOI	« Partially Depleted » Technologie Partiellement Désertée SOI
PMU	« Power Management Unit » Unité de gestion de puissance
RBB	« Reverse Body Biasing » Polarisation en inverse de body
R <sub>on</sub>	résistance des interrupteurs de puissance en mode passant et linéaire
R <sub>off</sub>	résistance des interrupteurs de puissance en mode bloqué
Setup	temps pré-positionnement
SOC	« System On Chip » Système sur puce
SOI	« Silicon-On-Insulator » Silicium sur isolant
STI	« Shallow Trench Isolation » Isolation latérale des transistors
SV <sub>t</sub>	tension de seuil standard
V <sub>D</sub>	tension de coude de la diode
VDD	tension d'alimentation
V <sub>drop</sub>	« Voltage drop » chute de tension aux bornes des interrupteurs de puissance
V <sub>th</sub> / V <sub>t</sub>	« Threshold Voltage » tension de seuil
V <sub>th0</sub>	tension de seuil sans polarisation
VTCMOS	« Variable Threshold CMOS » Transistor à tension de seuil variable





# Résumé

*Les technologies SOI partiellement désertées (PD-SOI) permettent de gagner en performance ou en consommation dynamique, par rapport à leur équivalent sur substrat massif (BULK). Leur inconvénient principal est la consommation statique qui est bien supérieure, en raison principalement de l'effet de body flottant de ses transistors.*

*Ce travail propose une technique de réduction de la consommation statique, pour la technologie PD-SOI, basée sur le principe des interrupteurs de puissance. Un nouveau facteur de mérite recherchant le meilleur compromis entre vitesse, courant de fuite et surface est introduit pour la sélection du meilleur interrupteur de puissance.*

*L'interrupteur de puissance proposé apporte par rapport à une solution de référence, et pour le même courant de fuite en mode éteint, une réduction de la résistance équivalente en mode passant de 20%.*

*Les tests comparatifs sur Silicium de blocs LDPC incluant ces montages montrent, par rapport au BULK, (1) un gain de 20% en vitesse pour la même tension d'alimentation ; (2) une réduction de 30% de la consommation dynamique pour la même vitesse ; (3) une division par 2 de la consommation statique.*

*Enfin, une bascule de rétention, élément nécessaire lorsque l'on met en œuvre des interrupteurs de puissance, est optimisée pour le PD-SOI. Cette bascule est conçue de façon à être robuste et peu fuyante.*

# Abstract

*Partially depleted SOI technologies (PD-SOI) exhibit advantages in terms of speed and dynamic power consumption, compared to bulk technologies. The main drawback of the PD-SOI technology is its static power consumption, which is higher than the bulk one. This comes from the floating body of its transistors.*

*This work presents a new static power consumption design technique based on power switches. A new figure of merit is introduced, allowing the selection of the power switch with the best trade-off in terms of leakage current, speed and area.*

*In addition, a new power switch architecture is introduced that brings, in comparison to a reference solution, a 20% reduction of the ON mode equivalent resistance for the same OFF mode leakage current.*

*This is validated on Silicon by supplying LDPC blocs with the proposed solution. Compared to the bulk technology, the following improvements were measured: (1) a 20% speed gain for the same supply voltage, (2) a 30% dynamic power consumption reduction at same speed; (3) a reduction of the static power consumption by a factor of 2.*

*Finally, a retention flip-flop, which is required when implementing power switches, is optimized in PD-SOI. This flip-flop is designed to be robust and exhibit a low leakage current.*